

GenCore version 5.1.4 p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 18, 2003, 04:46:26 ; Search time 278.727 Seconds
(without alignments)
2192.677 Million cell updates/sec

Title: US-09-270-437D-9
Perfect score: 21
Sequence: 1 gaaagtattcttaaggacgcc 21
Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.*

1: gb.ba.*

2: gb.htg.*

3: gb.in.*

4: gb.om.*

5: gb.ov.*

6: gb.pat.*

7: gb.ph.*

8: gb.pl.*

9: gb.pr.*

10: gb.ro.*

11: gb.sts.*

12: gb.sy.*

13: gb.un.*

14: gb.vi.*

15: em.ba.*

16: em.fun.*

17: em.hum.*

18: em.in.*

19: em.mu.*

20: em.om.*

21: em.or.*

22: em.ov.*

23: em.pat.*

24: em.ph.*

25: em.pl.*

26: em.ro.*

27: em.ste.*

28: em.un.*

29: em.vi.*

30: em.htg.hum.*

31: em.htg.inv.*

32: em.htg.other.*

33: em.htg.mus.*

34: em.htg.pln.*

35: em.htg.rod.*

36: em.htg.mam.*

37: em.htg.vrt.*

38: em.sv.*

39: em.htgo.hum.*

40: em.htgo.mus.*

41: em.htgo.other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	21	100.0	1740	6	AX365954	AX365954 Sequence
2	21	100.0	1743	6	AX366054	AX366054 Sequence
3	21	100.0	1743	6	AX366057	AX366057 Sequence
4	21	100.0	1764	6	AX366035	AX366035 Sequence
5	21	100.0	4155	9	HSU76705	U76705 Human putat
6	21	100.0	4159	6	ARI71863	ARI71863 Sequence
7	21	100.0	4171	9	AF117108	AF117108 Homo sapi
8	21	100.0	4181	6	AX333233	AX333233 Sequence
9	21	100.0	4181	6	AX365782	AX365782 Sequence
10	21	100.0	4181	6	HSU97188	U97188 Homo sapien
11	21	100.0	4602	6	AX397963	AX397963 Sequence
12	21	100.0	51119	2	AC090865	AC090865 Homo sapi
13	21	100.0	54379	2	AC079780	AC079780 Homo sapi
14	21	100.0	57261	2	AC090245	AC090245 Homo sapi
15	21	100.0	176547	9	AC092447	AC092447 Homo sapi
16	21	100.0	182312	2	AC087709	AC087709 Homo sapi
17	21	100.0	188389	2	AC024199	AC024199 Homo sapi
18	18.4	87.6	1740	10	AB046173	AB046173 Mus muscu
19	18.4	87.6	192406	2	AC115321	AC115321 Rattus no
20	17.8	84.8	60096	2	AC114540	AC114540 Mus muscu
21	17.8	84.8	109198	8	AC078891	AC078891 Oryza sat
22	17.8	84.8	129432	9	HS497321	AL023775 Human DNA
23	17.4	82.9	104983	2	AC119590	AC119590 Rattus no
24	17.4	82.9	167274	2	AC112310	AC112310 Rattus no
25	17.4	82.9	260605	2	AC099211	AC099211 Rattus no
26	17	81.0	143498	9	AC105391	AC105391 Homo sapi
27	17	81.0	159768	2	AC118170	AC118170 Rattus no
28	17	81.0	176348	2	AC102550	AC102550 Mus muscu
29	17	81.0	177435	2	AC128321	AC128321 Rattus no
30	17	81.0	194235	2	AC102586	AC102586 Mus muscu
31	17	81.0	201474	2	AC127286	AC127286 Mus muscu
32	16.8	80.0	2759	8	TGSUCSYN1	X96938 T.gesnerian
33	16.8	80.0	7661	10	AF012923	AF012923 Mus muscu
34	16.8	80.0	31170	3	CEF35E12	Z81527 Caenorhabdi
35	16.8	80.0	35016	9	AL603838	AL603838 Human DNA
36	16.8	80.0	81954	2	AC015133	AC015133 Drosophil
37	16.8	80.0	102477	2	AC110687	AC110687 Rattus no
38	16.8	80.0	114382	2	AC103498	AC103498 Rattus no
39	16.8	80.0	128750	2	AC007307	AC007307 Mus muscu
40	16.8	80.0	128750	2	AC007307	AC007307 Mus muscu
41	16.8	80.0	132171	3	AC008370	AC008370 Drosophil
42	16.8	80.0	133273	2	AC068586	AC068586 Homo sapi
43	16.8	80.0	133273	2	AC068586	AC068586 Homo sapi
44	16.8	80.0	140186	2	AC055808	AC055808 Homo sapi
45	16.8	80.0	157559	2	AC020344	AC020344 Drosophil

ALIGNMENTS

RESULT 1
AX365954
LOCUS
DEFINITION
Sequence 347 from Patent WO0200174.
ACCESSION
AX365954
VERSION
AX365954.1 GI:18697455
KEYWORDS
human.
SOURCE
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1
Wang, T., Wang, A., Skeiky, Y.A., Li, S.X., Kalos, M.D., Henderson, R.A.,
McNeill, P.D., Fanger, N., Retter, M.W., Marnerakis, M., Fanger, G.R.,
Vedvick, F.S., Carter, D., Watanabe, Y. and Peckham, D.W.

TITLE Compositions and methods for the therapy and diagnosis of lung

JOURNAL cancer
Patent: WO 0200174-A 347 03-JAN-2002;
CORIXA CORPORATION (US)

FEATURES Location/Qualifiers
source 1..1740

BASE COUNT 526 a 406 c 417 g 391 t
ORIGIN /db_xref="taxon:9606"

Query Match 100.0%; Score 21; DB 6; Length 1740;

Best Local Similarity 100.0%; Pred. No. 3.3; Indels 0; Gaps 0;
Matches 21; Conservative 0; Mismatches 0;

QY 1 GAAAGTATCTTCAAGGAGCC 21

Db 55 GAAAGTATCTTCAAGGAGCC 75

RESULT 2
AX366054 1743 bp DNA linear PAT 15-FEB-2002

LOCUS AX366054 Sequence 447 from Patent WO0200174.

DEFINITION AX366054

ACCESSION AX366054.1 GI:18697498

VERSION

KEYWORDS human.

SOURCE

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

AUTHORS Mammalia; Euthera; Primates; Catarrhini; Homnidae; Homo.

1 Wang, T., Wang, A., Skeiky, Y.A., Li, S.X., Kalos, M.D., Henderson, R.A.,

McNeill, P.D., Fanger, N., Retter, M.W., Marnerakis, M., Fanger, G.R.,

Vedvick, T.S., Carter, D., Watanabe, Y. and Peckham, D.W.

Compositions and methods for the therapy and diagnosis of lung

cancer

Patent: WO 0200174-A 447 03-JAN-2002;

CORIXA CORPORATION (US)

FEATURES Location/Qualifiers

source 1..1743

BASE COUNT 527 a 406 c 418 g 392 t

ORIGIN /db_xref="taxon:9606"

Query Match 100.0%; Score 21; DB 6; Length 1743;

Best Local Similarity 100.0%; Pred. No. 3.3;

Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAAAGTATCTTCAAGGAGCC 21

Db 55 GAAAGTATCTTCAAGGAGCC 75

RESULT 3

AX366057 1743 bp DNA linear PAT 15-FEB-2002

LOCUS AX366057 Sequence 450 from Patent WO0200174.

DEFINITION AX366057

ACCESSION AX366057.1 GI:18697500

VERSION

KEYWORDS human.

SOURCE

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

AUTHORS Mammalia; Euthera; Primates; Catarrhini; Homnidae; Homo.

1 Wang, T., Wang, A., Skeiky, Y.A., Li, S.X., Kalos, M.D., Henderson, R.A.,

McNeill, P.D., Fanger, N., Retter, M.W., Marnerakis, M., Fanger, G.R.,

Vedvick, T.S., Carter, D., Watanabe, Y. and Peckham, D.W.

Compositions and methods for the therapy and diagnosis of lung

cancer

Patent: WO 0200174-A 450 03-JAN-2002;

FEATURES Location/Qualifiers
source 1..1743

BASE COUNT 527 a 406 c 418 g 392 t
ORIGIN /db_xref="taxon:9606"

Query Match 100.0%; Score 21; DB 6; Length 1743;

Best Local Similarity 100.0%; Pred. No. 3.3;

Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAAAGTATCTTCAAGGAGCC 21

Db 55 GAAAGTATCTTCAAGGAGCC 75

RESULT 4

AX366035 1764 bp DNA linear PAT 15-FEB-2002

LOCUS AX366035 Sequence 428 from Patent WO0200174.

DEFINITION AX366035

ACCESSION AX366035.1 GI:18697484

VERSION

KEYWORDS human.

SOURCE

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

AUTHORS Mammalia; Euthera; Primates; Catarrhini; Homnidae; Homo.

1 Wang, T., Wang, A., Skeiky, Y.A., Li, S.X., Kalos, M.D., Henderson, R.A.,

McNeill, P.D., Fanger, N., Retter, M.W., Marnerakis, M., Fanger, G.R.,

Vedvick, T.S., Carter, D., Watanabe, Y. and Peckham, D.W.

Compositions and methods for the therapy and diagnosis of lung

cancer

Patent: WO 0200174-A 428 03-JAN-2002;

CORIXA CORPORATION (US)

FEATURES Location/Qualifiers

source 1..1764

BASE COUNT 534 a 417 c 419 g 394 t

ORIGIN /db_xref="taxon:9606"

Query Match 100.0%; Score 21; DB 6; Length 1764;

Best Local Similarity 100.0%; Pred. No. 3.3;

Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAAAGTATCTTCAAGGAGCC 21

Db 76 GAAAGTATCTTCAAGGAGCC 96

RESULT 5

HSU76705 4155 bp mRNA linear PRI 26-JAN-1999

LOCUS HSU76705 Human putative RNA binding protein Koc1 mRNA, complete cds.

DEFINITION U76705

ACCESSION U76705.1 GI:4098296

VERSION

KEYWORDS Homo sapiens.

SOURCE

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

AUTHORS Mammalia; Euthera; Primates; Catarrhini; Homnidae; Homo.

1 (bases 1 to 4155)

Mueller-Pillasch, F., Lacher, U. and Wallrapp, C.

Direct Submission

Submitted (30-OCT-1996) Innere Medizin I, University of Ulm,

Robert-Koch-Str. 8, Ulm 89081, Germany

JOURNAL

FEATURES Location/Qualifiers

source 1..4155

/db_xref="taxon:9606"

/chromosome="7"

CDs	/map="7p11.5"																						
	/cell_line="Patau8988s"																						
	/tissue_type="pancreas"																						
	251..1990																						
	/function="putative RNA binding protein"																						
	/notes="Kh-domain containing protein; overexpressed in cancer"																						
	/codon_start=1																						
	/product="Koc1"																						
	/protein_id="A009223.1"																						
	/db_xref="GI:4098297"																						
CDs	/translation="MNKLYIGNLSENAAPSDLESIFKDAKIPVSGPFLVKTGYAFVDC																						
	PDSWALKAIALSGKIELHGKPIEVHSPKQRIKQIRNIPPHLOWEVLDSL																						
	QYGVESCEQVNTDSETAVNVVYSSKQARQALDKNGFQLENFTLKVAYIPDEMAA																						
	QGNPLQQRGRGLQGRSGSGSKQPCDPLRLVPTQFVGAIIGKEGATI																						
	RNITKQTSKIDVHRKENAGAAEKSTITLSTPEGTSAAKSIILEIMHKEAQDIKFTTE																						
	IPLKILAHNFVGRLLIGKGRNLKIEQDITKITISPLQSLTYLNPFTTVKGNVE																						
	TCAAEEIMKKIRESYENDIASMNLQAHLLPGLNLAGLFPFTSGMPPPTSGPPSA																						
	MTPPYFQPSQSETVHQFIPALSVGAIIGKQGHILKLSRFAGASIKIAPAEAPDAK																						
	VRMVIITGPPEAQFKAQGRYIKIKEENFVSPKEEVKLEAHIRVPSFAAGRVIGKGGK																						
	TVNELQNLSSAEVVPRDQTPDENDQVVKITGHFYACQVQAKRIQEIILTQVKHQHQQ																						
BASE COUNT		1277 a	830 c	851 g	1181 t	16 others																	
ORIGIN																							
Query Match		100.0%; Score 21; DB 9; Length 4155;																					
Best Local Similarity		100.0%; Pred. No. 3.3;																					
Matches		21; Conservative	0; Mismatches	0; Indels	0; Gaps	0;																	
QY	Db	1 GAAAGTATCTTCAAGGACGCC 21																					
		305 GAAAGTATCTTCAAGGACGCC 325																					
RESULT 6																							
LOCUS		AR171863	4159 bp	DNA	linear	PAT 17-DEC-2001																	
DEFINITION		Sequence 4 from patent US 6297364.																					
ACCESSION		AR171863																					
VERSION		AR171863.1	GI:17910813																				
KEYWORDS																							
SOURCE		Unknown.																					
ORGANISM		Unknown.																					
REFERENCE		1 (bases 1 to 4159)																					
AUTHORS		Chen, Y.-T., Gure, A., Tsang, S., Stockert, E., Jager, E., Alexander, K. and Old, L.J.																					
TITLE		Isolated nucleic acid molecule encoding cancer associated antigen, the antigen itself, and uses thereof																					
JOURNAL		Patent: US 6297364-A 4 02-OCT-2001;																					
FEATURES		Location/Qualifiers																					
source		1..4159																					
BASE COUNT		1281 a	830 c	851 g	1181 t	16 others																	
ORIGIN																							
Query Match		100.0%; Score 21; DB 6; Length 4159;																					
Best Local Similarity		100.0%; Pred. No. 3.3;																					
Matches		21; Conservative	0; Mismatches	0; Indels	0; Gaps	0;																	
QY	Db	1 GAAAGTATCTTCAAGGACGCC 21																					
		305 GAAAGTATCTTCAAGGACGCC 325																					
RESULT 7																							
LOCUS		AF117108	4171 bp	mRNA	linear	PRI 26-JAN-1999																	
DEFINITION		Homo sapiens IGF-II mRNA-binding protein 3 (IMP-3) mRNA, complete cds.																					
ACCESSION		AF117108																					
VERSION		AF117108.1	GI:4191611																				
KEYWORDS																							
SOURCE		human.																					
ORGANISM		Homo sapiens																					
REFERENCE		1																					
AUTHORS		Young, P.E., Augustus, M., Carter, K.C., Ebner, R., Endress, G., Horigan, S., Soppet, D.R. and Weaver, Z.																					
TITLE		Cancer gene determination and therapeutic screening using signature gene sets																					
JOURNAL		Patent: WO 0194629-A 3742 13-DEC-2001;																					
BASE COUNT		1277 a	830 c	851 g	1181 t	16 others																	
ORIGIN																							
Query Match		100.0%; Score 21; DB 6; Length 4159;																					
Best Local Similarity		100.0%; Pred. No. 3.3;																					
Matches		21; Conservative	0; Mismatches	0; Indels	0; Gaps	0;																	
QY	Db	1 GAAAGTATCTTCAAGGACGCC 21																					
		305 GAAAGTATCTTCAAGGACGCC 325																					
RESULT 8																							
LOCUS		AX333233	4181 bp	DNA	linear	PAT 09-JAN-2002																	
DEFINITION		Sequence 3742 from Patent WO0194629.																					
ACCESSION		AX333233																					
VERSION		AX333233.1	GI:18123867																				
KEYWORDS																							
SOURCE		human.																					
ORGANISM		Homo sapiens																					
REFERENCE		1																					
AUTHORS		Young, P.E., Augustus, M., Carter, K.C., Ebner, R., Endress, G., Horigan, S., Soppet, D.R. and Weaver, Z.																					
TITLE		Cancer gene determination and therapeutic screening using signature gene sets																					
JOURNAL		Patent: WO 0194629-A 3742 13-DEC-2001;																					
BASE COUNT		1292 a	822 c	853 g	1199 t	5 others																	
ORIGIN																							
Query Match		100.0%; Score 21; DB 9; Length 4171;																					
Best Local Similarity		100.0%; Pred. No. 3.3;																					
Matches		21; Conservative	0; Mismatches	0; Indels	0; Gaps	0;																	
QY	Db	1 GAAAGTATCTTCAAGGACGCC 21																					
		321 GAAAGTATCTTCAAGGACGCC 341																					
RESULT 9																							
LOCUS		AX333233	4181 bp	DNA	linear	PAT 09-JAN-2002																	
DEFINITION		Sequence 3742 from Patent WO0194629.																					
ACCESSION		AX333233																					
VERSION		AX333233.1	GI:18123867																				
KEYWORDS																							
SOURCE		human.																					
ORGANISM		Homo sapiens																					
REFERENCE		1																					
AUTHORS		Young, P.E., Augustus, M., Carter, K.C., Ebner, R., Endress, G., Horigan, S., Soppet, D.R. and Weaver, Z.																					
TITLE		Cancer gene determination and therapeutic screening using signature gene sets																					
JOURNAL		Patent: WO 0194629-A 3742 13-DEC-2001;																					
BASE COUNT		1292 a	822 c	853 g	1199 t	5 others																	
ORIGIN																							
Query Match		100.0%; Score 21; DB 9; Length 4171;																					
Best Local Similarity		100.0%; Pred. No. 3.3;																					
Matches		21; Conservative	0; Mismatches	0; Indels	0; Gaps	0;																	
QY	Db	1 GAAAGTATCTTCAAGGACGCC 21																					
		321 GAAAGTATCTTCAAGGACGCC 341																					
RESULT 10																							
LOCUS		AX333233	4181 bp	DNA	linear	PAT 09-JAN-2002																	
DEFINITION		Sequence 3742 from Patent WO0194629.																					
ACCESSION		AX333233																					
VERSION		AX333233.1	GI:18123867																				
KEYWORDS																							
SOURCE		human.																					
ORGANISM		Homo sapiens																					
REFERENCE		1																					
AUTHORS		Young, P.E., Augustus, M., Carter, K.C., Ebner, R., Endress, G., Horigan, S., Soppet, D.R. and Weaver, Z.																					
TITLE		Cancer gene determination and therapeutic screening using signature gene sets																					
JOURNAL		Patent: WO 0194629-A 3742 13-DEC-2001;																					
BASE COUNT		1292 a	822 c	853 g	1199 t	5 others																	
ORIGIN																							
Query Match		100.0%; Score 21; DB 9; Length 4171;																					
Best Local Similarity		100.0%; Pred. No. 3.3;																					
Matches		21; Conservative	0; Mismatches	0; Indels	0; Gaps	0;																	
QY	Db	1 GAAAGTATCTTCAAGGACGCC 21																					
		321 GAAAGTATCTTCAAGGACGCC 341																					
RESULT 11																							
LOCUS		AX333233	4181 bp	DNA	linear	PAT 09-JAN-2002																	
DEFINITION		Sequence 3742 from Patent WO0194629.																					
ACCESSION		AX333233																					
VERSION		AX333233.1	GI:18123867																				
KEYWORDS																							
SOURCE		human.																					
ORGANISM		Homo sapiens																					
REFERENCE		1																					
AUTHORS		Young, P.E., Augustus, M., Carter, K.C., Ebner, R., Endress, G., Horigan, S., Soppet, D.R. and Weaver, Z.																					
TITLE		Cancer gene determination and therapeutic screening using signature gene sets																					
JOURNAL		Patent: WO 0194629-A 3742 13-DEC-2001;																					
BASE COUNT		1292 a	822 c	853 g	1199 t	5 others																	
ORIGIN																							
Query Match		100.0%; Score 21; DB 9; Length 4171;																					
Best Local Similarity		100.0%; Pred. No. 3.3;																					
Matches		21; Conservative	0; Mismatches	0; Indels	0; Gaps	0;																	
QY	Db	1 GAAAGTATCTTCAAGGACGCC 21																					
		321 GAAAGTATCTTCAAGGACGCC 341																					
RESULT 12																							
LOCUS		AX333233	4181 bp	DNA	linear	PAT 09-JAN-2002																	
DEFINITION		Sequence 3742 from Patent WO0194629.																					
ACCESSION		AX333233																					
VERSION		AX333233.1	GI:18123867																				
KEYWORDS																							
SOURCE		human.																					
ORGANISM		Homo sapiens																					
REFERENCE		1																					
AUTHORS		Young, P.E., Augustus, M., Carter, K.C., Ebner, R., Endress, G., Horigan, S., Soppet, D.R. and Weaver, Z.																					
TITLE		Cancer gene determination and therapeutic screening using signature gene sets																					
JOURNAL		Patent: WO 0194629-A 3742 13-DEC-2001;																					
BASE COUNT		1292 a	822 c	853 g	1199 t	5 others																	
ORIGIN																							
Query Match		100.0%; Score 21; DB 9; Length 4171;																					
Best Local Similarity		100.0%; Pred. No. 3.3;																					
Matches		21; Conservative	0; Mismatches	0; Indels	0; Gaps	0;																	
QY	Db	1 GAAAGTATCTTCAAGGACGCC 21																					
		321 GAAAGTATCTTCAAGGACGCC 341																					
RESULT 13																							
LOCUS		AX333233	4181 bp	DNA	linear	PAT 09-JAN-2002																	
DEFINITION		Sequence 3742 from Patent WO0194629.																					
ACCESSION		AX333233																					
VERSION		AX333233.1	GI:18123867																				
KEYWORDS																							
SOURCE		human.																					
ORGANISM		Homo sapiens																					
REFERENCE		1																					
AUTHORS		Young, P.E., Augustus, M., Carter, K.C., Ebner, R., Endress, G., Horigan, S., Soppet, D.R. and Weaver, Z.																					
TITLE		Cancer gene determination and therapeutic screening using signature gene sets																					
JOURNAL		Patent: WO 0194629-A 3742 13-DEC-2001;																					
BASE COUNT		1292 a	822 c	853 g	1199 t	5 others																	
ORIGIN																							
Query Match		100.0%; Score 21; DB 9; Length 4171;																					
Best Local Similarity		100.0%; Pred. No. 3.3;																					
Matches		21; Conservative	0; Mismatches	0; Indels	0; Gaps	0;																	
QY	Db	1 GAAAGTATCTTCAAGGACGCC 21																					
		321 GAAAGTATCTTCAAGGACGCC 341																					
RESULT 14																							
LOCUS		AX333233	4181 bp	DNA	linear	PAT 09-JAN-2002																	
DEFINITION		Sequence 3742 from Patent WO0194629.																					
ACCESSION		AX333233																					
VERSION		AX333233.1	GI:18123867																				
KEYWORDS																							
SOURCE		human.																					
ORGANISM		Homo sapiens																					
REFERENCE		1																					
AUTHORS		Young, P.E., Augustus, M., Carter, K.C., Ebner, R., Endress, G., Horigan, S., Soppet, D.R. and Weaver, Z.																					
TITLE		Cancer gene determination and therapeutic screening using signature gene sets																					
JOURNAL		Patent: WO 0194629-A 3742 13-DEC-2001;																					
BASE COUNT		1292 a	822 c	853 g	1199 t	5 others																	
ORIGIN																							
Query Match		100.0%; Score 21; DB 9; Length 4171;																					
Best Local Similarity		100.0%; Pred. No. 3.3;																					
Matches		21; Conservative	0; Mismatches	0; Indels	0; Gaps	0;																	
QY	Db	1 GAAAGTATCTTCAAGGACGCC 21																					
		321 GAAAGTATCTTCAAGGACGCC 341																					
RESULT 15																							
LOCUS		AX333233	4181 bp	DNA	linear	PAT 09-JAN-2002																	
DEFINITION		Sequence 3742 from Patent WO0194629.																					
ACCESSION		AX333233																					
VERSION		AX333233.1	GI:18123867																				
KEYWORDS																							
SOURCE		human.																					
ORGANISM		Homo sapiens																					
REFERENCE		1																					
AUTHORS		Young, P.E., Augustus, M., Carter, K.C., Ebner, R., Endress, G., Horigan, S., Soppet, D.R. and Weaver, Z.																					
TITLE		Cancer gene determination and therapeutic screening using signature gene sets																					
JOURNAL		Patent: WO 0194629-A 3742 13-DEC-2001;																					
BASE COUNT		1292 a	822 c	853 g	1199 t	5 others																	
ORIGIN																							
Query Match		100.0%; Score 21; DB 9; Length 4171;																					
Best Local Similarity		100.0%; Pred. No. 3.3;																					
Matches		21; Conservative	0; Mismatches	0; Indels	0; Gaps	0;																	
QY	Db	1 GAAAGTATCTTCAAGGACGCC 21																					
		321 GAAAGTATCTTCAAGGACGCC 341																					
RESULT 16																							
LOCUS		AX333233	4181 bp	DNA	linear	PAT 09-JAN-2002																	
DEFINITION		Sequence 3742 from Patent WO0194629.																					
ACCESSION		AX333233																					
VERSION		AX333233.1	GI:18123867																				
KEYWORDS																							
SOURCE		human.																					
ORGANISM		Homo sapiens																					
REFERENCE		1																					
AUTHORS		Young, P.E., Augustus, M., Carter, K.C., Ebner, R., Endress, G., Horigan, S., Soppet, D.R. and Weaver, Z.																					
TITLE		Cancer gene determination and therapeutic screening using signature gene sets																					
JOURNAL		Patent: WO 0194629-A 3742 13-DEC-2001;																					
BASE COUNT		1292 a	822 c	853 g	1199 t	5 others																	
ORIGIN																							
Query Match		100.0%; Score 21; DB 9; Length 4171;																					
Best Local Similarity		100.0%; Pred. No. 3.3;																					
Matches		21; Conservative	0; Mismatches	0; Indels	0; Gaps	0;																	
QY	Db	1 GAAAGTATCTTCAAGGACGCC 21																					
		321 GAAAGTATCTTCAAGGACGCC 341																					
RESULT 17																							
LOCUS		AX333233	4181 bp	DNA	linear	PAT 09-JAN-2002																	
DEFINITION		Sequence 3742 from Patent WO0194629.																					
ACCESSION		AX333233																					
VERSION		AX333233.1	GI:18123867																				
KEYWORDS																							
SOURCE		human.																					
ORGANISM		Homo sapiens																					
REFERENCE		1																					
AUTHORS		Young, P.E., Augustus, M., Carter, K.C., Ebner, R., Endress, G., Horigan, S., Soppet, D.R. and Weaver, Z.																					
TITLE		Cancer gene determination and therapeutic screening using signature gene sets																					
JOURNAL		Patent: WO 0194629-A 3742 13-DEC-2001;																					
BASE COUNT		1292 a	822 c	853 g	1199 t	5 others																	
ORIGIN																							
Query Match		100.0%; Score 21; DB 9; Length 4171;																					
Best Local Similarity		100.0%; Pred. No. 3.3;																					
Matches		21; Conservative	0; Mismatches	0; Indels	0; Gaps	0;																	
QY	Db	1 GAAAGTATCTTCAAGGACGCC 21																					
		321 GAAAGTATCTTCAAGGACGCC 341																					
RESULT 18																							
LOCUS		AX333233	4181 bp	DNA	linear	PAT 09-JAN-2002																	
DEFINITION		Sequence 3742 from Patent WO0194629.																					
ACCESSION		AX333233																					
VERSION		AX333233.1	GI:18123867																				
KEYWORDS																							
SOURCE		human.																					
ORGANISM		Homo sapiens																					
REFERENCE		1																					
AUTHORS		Young, P.E., Augustus, M., Carter, K.C., Ebner, R., Endress, G., Horigan, S., Soppet, D.R. and Weaver, Z.																					
TITLE		Cancer gene determination and therapeutic screening using signature gene sets																					
JOURNAL		Patent: WO 0194629-A 3742 13-DEC-2001;																					
BASE COUNT		1292 a	822 c	853 g	1199 t	5 others																	
ORIGIN																							
Query Match		100.0%; Score 21; DB 9; Length 4171;																					
Best Local Similarity		100.0%; Pred. No. 3.3;																					
Matches		21; Conservative	0; Mismatches	0; Indels	0; Gaps	0;																	
QY	Db	1 GAAAGTATCTTCAAGGACGCC 21																					
		321 GAAAGTATCTTCAAGGACGCC 341																					

```
6/2: .Z411
/note="unnamed protein product"
```



```

/codon_start=1
/db_xref="GI:21260819"
/translation="MNKLYIGNLSENAAPSDLESIFDKAKIPVSGPLVTGTGAFVDC
PDESWALKAEALSGKIEHGKIEVSHVPRQRIRKLQIRNIPPHLQWELVSLV
OYGVSECEQVNTDSEAVNTVYSSDKQARQALDKLNGFQLENFTLKVAIIPDEMAA
QNFPLQQRGRGGLGQSSGSGVSKQKCDPLRLVLPVTFQVGAIGKEGATI
ITPKIHLANNFVGRIGKEGNLKKIBQDITDTKITISPLQELTYNPERITVKGNE
TCAKAEIIMKIRSEYENDIASMNLQHLIPGLNLAGLIPPTSMPPPTSGPSA
MTPYVPOFEQSETVTHQFIPALSVGAILGKQGHKQLSRFAGASTIKIAPAPDAK
VRVLIITGPPEAQKAGRIYKKEENFVSPKEEVLEAHIRVPSFAAGRVTKGCK
TVNQLNLSRAEVPRDQTPDENQVVKITGHFYACQVQAKRIQBIILTQVQHQQK
KALQSGPPQSRKK"
BASE COUNT      1463 a      907 c      918 g      1298 t      16 others
ORIGIN
Query Match      100.0%; Score 21; DB 6; Length 4602;
Best Local Similarity 100.0%; Pred. NO. 3.4;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Oy 1 GAAAGTATCTTCAAGGACGCC 21
|||||
Db 726 GAAAGTATCTTCAAGGACGCC 746

RESULT 12
AC090865
LOCUS
DEFINITION      AC090865      51119 bp      DNA      linear      HTG 13-APR-2001
SEQUENCE SAMPLING.
AC090865
VERSION
KEYWORDS
SOURCE
ORGANISM
Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 51119)
Birren,B., Linton,L., Nuebaum,C., Lander,E., Allen,N., Anderson,S.,
Camarata,J., Campopiano,A., Choepel,Y., Colangelo,M., Collins,S.,
Collamore,A., Cooke,P., DeArrellano,K., Dewar,K., Diaz,J.S.,
Dodge,S., Faro,S., Ferrelira,P., FitzHugh,W., Gage,D., Galagan,J.,
Gardyna,S., Ginde,S., Govette,M., Graham,L., Grand-Pierre,N.,
Hagos,B., Heaford,A., Horton,L., Hulme,W., Iliev,I., Johnson,R.,
Jones,C., Karatas,A., LaRocque,K., Lamazares,R., Landers,T.,
Lehoczy,J., Levine,R., Liu,G., MacLean,C., Macdonald,P.,
Marquis,N., Matthews,C., McCarthy,M., McEwan,P., McKernan,K.,
McPheeters,R., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V.,
Murphy,T., Naylor,J., Nguyen,C., Norbu,C., Norman,C.H.,
O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J., Peterson,K.,
Phunkhang,P., Pierre,N., Pollara,V., Raymond,C., Retta,R.,
Rieback,M., Riley,R., Rise,C., Rogov,P., Roman,J., Rosetti,M.,
Roy,A., Santos,R., Schauer,S., Schuback,R., Seaman,S., Severy,P.,
Souniez,C., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
Travers,M., Travis,N., Triglio,J., Vassiliev,H., Viel,R., Vo,A.,
Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J.,
Zembek,L., Zimmer,A. and Zody,M.
Direct Submission
Submitted (11-MAR-2001) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Apr 13, 2001 this sequence version replaced gi:13273446.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html.
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research

```

Center code: WIBR
 Web site: <http://www-seq.wi.mit.edu>
 Contact: sequence_submissions@genome.wi.mit.edu
 ----- Project Information
 Center project name: L13138
 Center clone name: 494_F_2

* NOTE: This record contains 61 individual
 * sequencing reads that have not been assembled into
 * contigs. Runs of N are used to separate the reads
 * and the order in which they appear is completely
 * arbitrary. Low-pass sequence sampling is useful for
 * identifying clones that may be gene-rich and allows
 * overlap relationships among clones to be deduced.
 * However, it should not be assumed that this clone
 * will be sequenced to completion. In the event that
 * the record is updated, the accession number will
 * be preserved.

* 1 768: contig of 768 bp in length
 * 769 868: gap of 100 bp
 * 869 1597: contig of 729 bp in length
 * 1598 1697: gap of 100 bp
 * 1698 2452: contig of 755 bp in length
 * 2453 2552: gap of 100 bp
 * 2553 3294: contig of 742 bp in length
 * 3295 3394: gap of 100 bp
 * 3395 4151: contig of 757 bp in length
 * 4152 4551: gap of 100 bp
 * 4552 4995: contig of 744 bp in length
 * 4996 5095: gap of 100 bp
 * 5096 5848: contig of 753 bp in length
 * 5849 5948: gap of 100 bp
 * 5949 6689: contig of 741 bp in length
 * 6690 6789: gap of 100 bp
 * 6790 7531: contig of 742 bp in length
 * 7532 7631: gap of 100 bp
 * 7632 8361: contig of 730 bp in length
 * 8362 8461: gap of 100 bp
 * 8462 9202: contig of 741 bp in length
 * 9203 9302: gap of 100 bp
 * 9303 10050: contig of 748 bp in length
 * 10051 10150: gap of 100 bp
 * 10151 10868: contig of 718 bp in length
 * 10869 10968: gap of 100 bp
 * 10969 11718: contig of 750 bp in length
 * 11719 11818: gap of 100 bp
 * 11819 12541: contig of 723 bp in length
 * 12542 12641: gap of 100 bp
 * 12642 13349: contig of 708 bp in length
 * 13350 13449: gap of 100 bp
 * 13450 14174: contig of 725 bp in length
 * 14175 14274: gap of 100 bp
 * 14275 14996: contig of 722 bp in length
 * 14997 15096: gap of 100 bp
 * 15097 15827: contig of 731 bp in length
 * 15828 15927: gap of 100 bp
 * 15928 16677: contig of 750 bp in length
 * 16678 16777: gap of 100 bp
 * 16778 17496: contig of 719 bp in length
 * 17497 17596: gap of 100 bp
 * 17597 18342: contig of 746 bp in length
 * 18343 18442: gap of 100 bp
 * 18443 19193: contig of 751 bp in length
 * 19194 19293: gap of 100 bp
 * 19294 20044: contig of 751 bp in length
 * 20045 20444: gap of 100 bp
 * 20445 20872: contig of 728 bp in length
 * 20873 20972: gap of 100 bp
 * 20973 21715: contig of 743 bp in length
 * 21716 21815: gap of 100 bp
 * 21816 22547: contig of 732 bp in length
 * 22548 22647: gap of 100 bp
 * 22648 23401: contig of 754 bp in length

23402 23501: gap of 100 bp
23502 24255: contig of 754 bp in length
24256 24355: gap of 100 bp
24356 25124: contig of 769 bp in length
25125 25224: gap of 100 bp
25225 25962: contig of 738 bp in length
25963 26062: gap of 100 bp
26063 26786: contig of 724 bp in length
26787 26886: gap of 100 bp
26887 27593: contig of 707 bp in length
27594 27693: gap of 100 bp
27694 28439: contig of 746 bp in length
28440 28539: gap of 100 bp
28540 29277: contig of 738 bp in length
29278 29377: gap of 100 bp
29378 30137: contig of 760 bp in length
30138 30237: gap of 100 bp
30238 30968: contig of 731 bp in length
30969 31068: gap of 100 bp
31069 31795: contig of 727 bp in length
31796 31895: gap of 100 bp
31896 32638: contig of 743 bp in length
32639 32738: gap of 100 bp
32739 33497: contig of 759 bp in length
33498 33597: gap of 100 bp
33598 34339: contig of 742 bp in length
34340 34439: gap of 100 bp
34440 35167: contig of 728 bp in length
35168 35267: gap of 100 bp
35268 36015: contig of 748 bp in length
36016 36115: gap of 100 bp
36116 36878: contig of 763 bp in length
36879 36978: gap of 100 bp
36979 37744: contig of 766 bp in length
37745 37844: gap of 100 bp
37845 38596: contig of 752 bp in length
38597 38696: gap of 100 bp
38697 39433: contig of 737 bp in length
39434 39533: gap of 100 bp
39534 40286: contig of 753 bp in length
40287 40386: gap of 100 bp
40387 41064: contig of 678 bp in length
41065 41164: gap of 100 bp
41165 41895: contig of 731 bp in length
41896 41995: gap of 100 bp
41996 42740: contig of 745 bp in length
42741 42840: gap of 100 bp
42841 43593: contig of 753 bp in length
43594 43693: gap of 100 bp
43694 44422: contig of 729 bp in length
44423 44522: gap of 100 bp
44523 45270: contig of 748 bp in length
45271 45370: gap of 100 bp
45371 46108: contig of 738 bp in length
46109 46208: gap of 100 bp
46209 46923: contig of 715 bp in length
46924 47023: gap of 100 bp
47024 47753: contig of 730 bp in length
47754 47853: gap of 100 bp
47854 48595: contig of 742 bp in length
48596 48695: gap of 100 bp
48696 49432: contig of 737 bp in length
49433 49532: gap of 100 bp
49533 50267: contig of 735 bp in length
50268 50367: gap of 100 bp
50368 51119: contig of 752 bp in length.

FEATURES
Source

1. 51119
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="15"
/map="15"
/clone="RP11-494F2"

BASE COUNT 12444 a 10521 c 9576 g 11843 t 6735 others
ORIGIN
Query Match 100.0%; Score 21; DB 2; Length 51119;
Best Local Similarity 100.0%; Pred. No. 3.5; Indels 0; Gaps 0;
Matches 21; Conservative 0; Mismatches 0
Qy 1 GAAAGTATCTTCAAGGACGCC 21
|||||
Db 28678 GAAAGTATCTTCAAGGACGCC 28698
|||||
RESULT 13
AC079780/c 54379 bp DNA linear PRI 21-FEB-2002
LOCUS Homo sapiens BAC clone RP11-379L16 from 7, complete sequence.
DEFINITION AC079780
ACCESSION AC079780.6 GI:18072220
VERSION HTG.
KEYWORDS
SOURCE Homo sapiens.
ORGANISM
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 54379)
AUTHORS Sulston, J.E. and Waterston, R.
TITLE Toward a complete human genome sequence
JOURNAL Genome Res. 8 (11), 1097-1108 (1998)
MEDLINE 99063792
PUBMED 9847074
REFERENCE 2 (bases 1 to 54379)
AUTHORS Swearingen, S., Meyer, R. and Cotton, M.
TITLE The sequence of Homo sapiens BAC clone RP11-379L16
JOURNAL Unpublished (2001)
REFERENCE 3 (bases 1 to 54379)
AUTHORS Waterston, R.H.
TITLE Direct Submission
JOURNAL Submitted (10-SEP-2000) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
REFERENCE 4 (bases 1 to 54379)
AUTHORS Waterston, R.H.
TITLE Direct Submission
JOURNAL Submitted (06-JAN-2002) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
REFERENCE 5 (bases 1 to 54379)
AUTHORS Waterston, R.
TITLE Direct Submission
JOURNAL Submitted (21-FEB-2002) Department of Genetics, Washington
University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
On Jan 6, 2002 this sequence version replaced gi:17298647.

Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: <http://genome.wustl.edu/gsc>
Contact: sapiens@wustl.edu

Summary Statistics

Center project name: H_NH0379L16

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted:
all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:

The sequence of this clone was established as part of a mapping and sequencing collaboration between the NHGRI Chromosome 7 Mapping Project (Eric D. Green, Director), John D. McPherson in the Department of Genetics (Washington University), and the Washington University Genome Sequencing Center. For additional information about the map position of this sequence, see <http://www.nhgri.nih.gov/DIR/STB/CHR7>, send <mailto:regreen@nhgri.nih.gov>, or see <http://genome.wustl.edu/gsc>

SOURCE INFORMATION:

The RPCI-11 human BAC library was made from the blood of one male donor, as described by Osoegawa, K., Woon, P. Y., Zhao, B., Frangen, E., Tateno, M., Catanese, J. J., and de Jong, P. J. (1998) An improved approach for construction of bacterial artificial chromosome libraries. *Genomics* 51:1-8. The clone may be obtained either from Research Genetics, Inc. (<http://www.resgen.com>) or Pieter de Jong and coworkers at <http://www.choiri.org> VECTOR: pBACe3.6

NEIGHBORING SEQUENCE INFORMATION:

The clone sequenced to the left is GS1-117B4, 2000 bp overlap; the clone sequenced to the right is GS1-13M19, 2000 bp overlap. Actual start of this clone is at base position 25835 of GS1-117B4; actual end is at base position 49313 of GS1-13M19.

Polymorphisms have been identified between AC023105 and AC079780.

FEATURES

source

1. .54379
/location="Qualifiers"
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="7"
/map="7"

/clone="RP11-379L16"
/clone_lib="RPCI-11"

repeat_region

1. .70

repeat_region

72..370

repeat_region

382..809

misc_feature

795..1200

repeat_region

1150..1454

repeat_region

1516..1828

repeat_region

1802..1840

repeat_region

2087..2289

repeat_region

2447..2725

repeat_region

2895..2939

repeat_region

2927..2949

repeat_region

3008..3054

repeat_region

3280..3575

repeat_region

3554..3575

repeat_region

3587..3893

repeat_region

3871..3893

repeat_region

4308..4602

repeat_region

4603..4759

repeat_region

5152..5722

misc_feature

misc_feature
5335..5644
/notes="match to EST BF825880 (NID:gl12168647)"
repeat_region
5515..5592
/notes="match to EST BE149137 (NID:g8611874)"
repeat_region
5609..5920
/rpt_family="Alu"
misc_feature
5780..5810
/rpt_family="Alu"
repeat_region
5891..5920
/notes="match to EST BE149137 (NID:g8611874)"
repeat_region
5911..5920
/rpt_family="(A)n"
repeat_region
6159..6450
/rpt_family="Alu"
repeat_region
6502..6566
/rpt_family="L2"
repeat_region
6714..7011
/rpt_family="Alu"
repeat_region
7419..7446
/rpt_family="(T)n"
repeat_region
7742..7781
/rpt_family="Alu"
repeat_region
7782..8020
/rpt_family="Alu"
repeat_region
7989..8031
/rpt_family="(A)n"
repeat_region
8044..8351
/rpt_family="AT_rich"
repeat_region
8775..9081
/rpt_family="Alu"
repeat_region
9059..9081
/rpt_family="(A)n"
repeat_region
9108..9390
/rpt_family="Alu"
misc_feature
10078..10478
/notes="match to EST AA112671 (NID:gl1665389) zn69h10.r1"
repeat_region
10379..10454
/rpt_family="L2"
repeat_region
10461..10592
/rpt_family="L1"
repeat_region
10593..10898
/rpt_family="Alu"
repeat_region
10899..11237
/rpt_family="L1"
repeat_region
11238..11512
/rpt_family="Alu"
repeat_region
11513..11668
/rpt_family="L1"
repeat_region
11669..11965
/rpt_family="Alu"
repeat_region
11966..12004
/rpt_family="L1"
repeat_region
12007..12320
/rpt_family="Alu"
repeat_region
12292..12312
/rpt_family="AT_rich"
repeat_region
12739..12937
/rpt_family="Alu"
repeat_region
12932..12965
/rpt_family="(TTTG)n"
repeat_region
12938..13246
/rpt_family="Alu"
repeat_region
13247..13341
/rpt_family="Alu"
misc_feature
13448..13617
/notes="match to EST AA301911 (NID:gl1954244)"
repeat_region
13552..13836
/rpt_family="Alu"
repeat_region
13849..13892
/rpt_family="(T)n"

Query Match

Best Local Similarity .100.0%; Score 21; DB 9; Length 54379;

Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Mon Apr 21 10:30:47 2003

```

OY      1  GAAAGTATCTTCAAGACGCC 21
DB 34372 GAAAGTATCTTCAAGACGCC 34352

RESULT 14
AC090245/C
LOCUS   Homo sapiens chromosome 18 clone RP11-888I12 map 18, LOW-PASS
DEFINITION
SEQUENCE SAMPLING.
ACCESSION AC090245
VERSION   AC090245.2 GI:13374670
KEYWORDS  HTG; HTGS_PHASE0.
SOURCE    Homo sapiens.
ORGANISM  Homo sapiens
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
  1 (bases 1 to 57261)
  Birren, B., Linton, L., Nusbaum, C., Lander, E., Allen, N., Anderson, S.,
  Barna, N., Bastien, V., Boguslavskiy, L., Boukangelter, B., Brown, A.,
  Camarata, J., Campopiano, A., Choepel, Y., Colangelo, M., Collins, S.,
  Collymore, A., Cooke, P., DeArellano, K., Dewar, K., Diaz, J.S.,
  Dodge, S., Faro, S., Ferreira, P., FitzHugh, W., Gage, D., Galagan, J.,
  Gardyna, S., Ginde, S., Govette, M., Graham, L., Grand-Pierre, N.,
  Hagos, B., Heaford, A., Horton, L., Hulme, W., Iliev, I., Johnson, R.,
  Jones, C., Karatas, A., Larkocque, K., Lamazares, R., Landers, T.,
  Lehoczy, J., Levine, R., Liu, G., MacLean, C., Macdonald, P.,
  Marquis, N., Matthews, C., McCarthy, M., McEwan, P., McKernan, K.,
  McPheters, R., Meldrim, J., Meneus, L., Mihova, T., Mlenga, V.,
  Murphy, T., Naylor, J., Nguyen, C., Norbu, C., Norman, C.H.,
  O'Connor, T., O'Donnell, P., O'Neil, D., Oliver, J., Peterson, K.,
  Phunkhang, P., Pierre, N., Pollara, V., Raymond, C., Retta, R.,
  Rieback, M., Riley, R., Rise, C., Rogov, P., Roman, J., Rosetti, M.,
  Roy, A., Santos, R., Schauer, S., Schupbach, R., Seaman, S., Severy, P.,
  Sounez, C., Spencer, B., Stange-Thomann, N., Stojanovic, N.,
  Strauss, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J.,
  Travers, M., Travis, N., Trigilio, J., Vassiliev, H., Viel, R., Vo, A.,
  Wilson, B., Wu, X., Wyman, D., Ye, W.J., Young, G., Zainoun, J.,
  Zembek, L., Zimmer, A. and Zody, M.
Direct Submission
Submitted (17-FEB-2001) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Mar 17, 2001 this sequence version replaced gi:12957879.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence.submissions@genome.wi.mit.edu
----- Project Information
Center project name: L12326
Center clone name: 888_I_12
-----
* NOTE: This record contains 71 individual
* sequencing reads that have not been assembled into
* contigs. Runs of N are used to separate the reads
* and the order in which they appear is completely
* arbitrary. Low-pass sequence sampling is useful for
* identifying clones that may be gene-rich and allows
* overlap relationships among clones to be deduced.
* However, it should not be assumed that this clone
* will be sequenced to completion. In the event that
* the record is updated, the accession number will
* be preserved.
*
* 1
* 718 817: gap of 100 bp
* 717: contig of 717 bp in length

```

```

* 30633 30732: contig of 100 bp
* 30733 31440: contig of 708 bp in length
* 31441 31540: gap of 100 bp
* 31541 32356: contig of 716 bp in length
* 32357 32356: gap of 100 bp
* 32357 33051: contig of 695 bp in length
* 33052 33151: gap of 100 bp
* 33152 33864: contig of 713 bp in length
* 33865 33964: gap of 100 bp
* 33965 34677: contig of 713 bp in length
* 34678 34777: gap of 100 bp
* 34778 35494: contig of 717 bp in length
* 35495 35594: gap of 100 bp
* 35595 36312: contig of 718 bp in length
* 36313 36412: gap of 100 bp
* 36413 37086: contig of 674 bp in length
* 37087 37186: gap of 100 bp
* 37187 37877: contig of 691 bp in length
* 37878 37977: gap of 100 bp
* 37978 38681: contig of 704 bp in length
* 38682 38781: gap of 100 bp
* 38782 39450: contig of 669 bp in length
* 39451 39550: gap of 100 bp
* 39551 40238: contig of 688 bp in length
* 40239 40338: gap of 100 bp
* 40339 41055: contig of 717 bp in length
* 41056 41155: gap of 100 bp
* 41156 41873: contig of 718 bp in length
* 41874 41973: gap of 100 bp
* 41974 42694: contig of 721 bp in length
* 42695 42794: gap of 100 bp
* 42795 43495: contig of 701 bp in length
* 43496 43595: gap of 100 bp
* 43596 44320: contig of 725 bp in length
* 44321 44420: gap of 100 bp
* 44421 45137: contig of 707 bp in length
* 45128 45227: gap of 100 bp
* 45228 45922: contig of 705 bp in length
* 45933 46032: gap of 100 bp
* 46033 46734: contig of 702 bp in length
* 46735 46834: gap of 100 bp
* 46835 47535: contig of 701 bp in length
* 47536 47635: gap of 100 bp
* 47636 48346: contig of 711 bp in length
* 48347 48446: gap of 100 bp
* 48447 49130: contig of 684 bp in length
* 49131 49230: gap of 100 bp
* 49231 49954: contig of 724 bp in length
* 49955 50054: gap of 100 bp
* 50055 50770: contig of 716 bp in length
* 50771 50870: gap of 100 bp
* 50871 51594: contig of 724 bp in length
* 51595 51694: gap of 100 bp
* 51695 52402: contig of 708 bp in length
* 52403 52502: gap of 100 bp
* 52503 53213: contig of 711 bp in length
* 53214 53313: gap of 100 bp
* 53314 54026: contig of 713 bp in length
* 54027 54126: gap of 100 bp
* 54127 54831: contig of 705 bp in length
* 54832 54931: gap of 100 bp
* 54932 55624: contig of 693 bp in length

```

```

Query Match      100.0%; Score 21; DB 2; Length 57261;
Best Local Similarity 100.0%; Pred. No. 3.6;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 GAAAGTATCTTCAAGGACCC 21

```

```

DB 28686 GAAAGTATCTTCAAGGACCC 28666

```

```

RESULT 15

```

```

AC092447
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL

```

```

AC092447
Homo sapiens BAC clone RP11-760D2 from 7, complete sequence.
AC092447
AC092447.5 GI:18873962
HTG
Homo sapiens.
Homo sapiens.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 176547)
Sulston, J.B. and Waterston, R.
Toward a complete human genome sequence
Genome Res. 8 (11), 1097-1108 (1998)
99063792
9847074
2 (bases 1 to 176547)
Armstrong, J. and Cotton, M.
The sequence of Homo sapiens BAC clone RP11-760D2
Unpublished (2001)
3 (bases 1 to 176547)
Waterston, R.H.
Direct Submission
Submitted (04-JUL-2001) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
4 (bases 1 to 176547)
Waterston, R.H.
Direct Submission
Submitted (23-FEB-2002) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
5 (bases 1 to 176547)
Waterston, R.H.
Direct Submission
Submitted (27-FEB-2002) Department of Genetics, Washington
University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
On Feb 23, 2002 this sequence version replaced gi:18259201.
----- Genome Center
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu/gsc
Contact: sapiens@wustl.wustl.edu
----- Summary Statistics
Center project name: H_NH0760D02
-----

```

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted: all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:

The sequence of this clone was established as part of a mapping and sequencing collaboration between the NHGRI Chromosome 7 Mapping Project (Eric D. Green, Director), John D. McPherson in the Department of Genetics (Washington University), and the Washington University Genome Sequencing Center. For additional information about the map position of this sequence, see <http://www.nhgri.nih.gov/DIR/GRB/CHR7>, send mail to: egreen@nhgri.nih.gov, or see <http://genome.wustl.edu/gsc>

SOURCE INFORMATION:

The RPCI-11 human BAC library was made from the blood of one male donor, as described by Osoegawa, K., Woon, P.Y., Zhao, B., Frengen, E.,

Tateno, M., Catanese, J. J. and de Jong, P. J. (1998) An improved approach for construction of bacterial artificial chromosome libraries. Genomics 51:1-8. The clone may be obtained either from Research Genetics, Inc. (<http://www.resgen.com>) or Pieter de Jong and coworkers at <http://www.chori.org>
VECTOR: pBACE3.6

NEIGHBORING SEQUENCE INFORMATION:
The clone sequenced to the left is RP13-492C18; the clone sequenced to the right is RP11-10F11. 2000 bp overlap. Actual start of this clone is at base position 1 of RP11-760D2; actual end is at base position 17615 of RP11-10F11.

The region from 20341 to 20385 is covered only by a PCR product from clone DNA. An unresolved simple sequence repeats exists between 84203 and 84724. Polymorphisms exists between AC024199 and AC092447. Data from AC024199 and AC092423 was used to finish AC092447.

FEATURES

source

Location/Qualifiers
1. .176547
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="7"
/map="7"
/clone="RP11-760D2"
/clone_lib="RPCI-11"

repeat_region

5. .1938
/rpt_family="L1"
2210. .2483
/rpt_family="Alu"

repeat_region

2799. .3012
/note="match to EST BF994500 (NID:g12285959)"

misc_feature

2799. .2962
/note="match to EST BF995059 (NID:g12286518)"

repeat_region

3092. .3388
/rpt_family="L2"
3419. .3536
/rpt_family="L1"

repeat_region

3537. .3794
/rpt_family="Alu"

repeat_region

3795. .4263
/rpt_family="L1"

repeat_region

4259. .4542
/rpt_family="L1"

repeat_region

4548. .4748
/rpt_family="L1"

repeat_region

4749. .5046
/rpt_family="Alu"

repeat_region

5047. .5267
/rpt_family="L1"

repeat_region

5268. .5560
/rpt_family="Alu"

repeat_region

5561. .6091
/rpt_family="L1"

repeat_region

6528. .6825
/rpt_family="Alu"

repeat_region

8380. .8686
/rpt_family="Alu"

repeat_region

8924. .9225
/rpt_family="Alu"

repeat_region

9838. .10125
/rpt_family="Alu"

repeat_region

10336. .10600
/rpt_family="Alu"

repeat_region

11975. .12282
/rpt_family="Alu"

repeat_region

12457. .12765
/rpt_family="Alu"

misc_feature

14411. .14448
/note="match to EST AI334287 (NID:g4070846) qq25e08.xl"

misc_feature

14411. .14446
/note="similar to Homo sapiens EST BF979990 (NID:g12347205)"

misc_feature

14412. .14455
/note="similar to Homo sapiens EST BE879245 (NID:g10328021)"

misc_feature

14412. .14448
/note="similar to Homo sapiens EST AL520426 (NID:g12783919)"

repeat_region

14530. .14608
/rpt_family="L1"

repeat_region

14609. .14910
/rpt_family="Alu"

repeat_region

14911. .15079
/rpt_family="L1"

repeat_region

15062. .15811
/rpt_family="L1"

repeat_region

15812. .16112
/rpt_family="Alu"

repeat_region

16113. .16529
/rpt_family="L1"

misc_feature

16530. .16941
/note="similar to Homo sapiens EST AA446576 (NID:g2159241) zw84d02.81"

misc_feature

16541. .17068
/note="similar to Homo sapiens EST BE966979 (NID:g11772953)"

misc_feature

16586. .17133
/note="similar to Homo sapiens EST BE622588 (NID:g9893515)"

misc_feature

16610. .17149
/note="similar to Homo sapiens EST AI744560 (NID:g5112848) wg09g02.x1"

misc_feature

16743. .17248
/note="similar to Homo sapiens EST N46675 (NID:g1187841) yy50c07.r1"

misc_feature

16912. .17332
/note="similar to Homo sapiens EST AA446703 (NID:g2159368) zw84d02.r1"

misc_feature

16921. .17383
/note="match to EST BG479148 (NID:g13411427)"

repeat_region

17135. .17383
/note="match to EST BG773852 (NID:g14044151)"

repeat_region

17144. .17383
/note="similar to Homo sapiens EST BG480641 (NID:g13412920)"

repeat_region

17383. .17822
/rpt_family="L1"

repeat_region

17826. .18132
/rpt_family="Alu"

misc_feature

17826. .17839
/note="similar to Homo sapiens EST AA669082 (NID:g2630581) ab92h09.81"

repeat_region

18142. .18846
/rpt_family="ERV1"

misc_feature

18530. .18619
/note="match to EST BF894500 (NID:g12285959)"

repeat_region

19083. .19168
/note="match to EST BF894500 (NID:g12285959)"

Query Match

Best Local Similarity 100.0%; Score 21; DB 9; Length 176547;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY

1 GAAAGTATCTTCAAGGACGCC 21
|||||
Db 167744 GAAAGTATCTTCAAGGACGCC 167764

Search completed:

April 18, 2003, 06:15:49
Job time : 345.727 secs

GenCore version 5.1.4.DS 4578.
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 18, 2003, 04:12:35 ; Search time 93.5455 Seconds
(without alignments)
505.551 Million cell updates/sec

Title: US-09-270-437D-9
Perfect score: 21
Sequence: 1 gaagatcttcaaggagcc 21
Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_101002.*

- 1: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*
- 2: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*
- 3: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*
- 4: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*
- 5: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*
- 6: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*
- 7: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*
- 8: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*
- 9: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*
- 10: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*
- 11: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*
- 12: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*
- 13: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*
- 14: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*
- 15: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*
- 16: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*
- 17: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*
- 18: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*
- 19: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*
- 20: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*
- 21: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*
- 22: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*
- 23: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*
- 24: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	21	100.0	21	21	AAZ36155
2	21	100.0	1740	21	AAZ36155
3	21	100.0	1740	24	ABL49254
4	21	100.0	1743	24	ABL49297
5	21	100.0	1743	24	ABL49299
6	21	100.0	1764	24	ABL49283
7	21	100.0	4159	21	AAZ36150
8	21	100.0	4181	21	AAZ36150
9	21	100.0	4181	24	ABL65405

10	21	100.0	4181	24	ABL49119	Human lung tumour
11	21	100.0	4601	24	ABA99958	Human KOC DNA. Ho
12	20	95.2	4264	23	AAZ36150	DNA encoding novel
13	17.8	84.8	1985	23	AAZ36150	DNA encoding novel
14	16.8	80.0	7661	24	ABO88221	Human osteoblast d
15	16.4	78.1	557	23	ABL13727	Drosophila melanog
16	16.4	78.1	1421	23	ABL20853	Drosophila melanog
17	16.4	78.1	1920	22	AAH14309	Human cDNA sequenc
18	16.4	78.1	2557	23	ABL13726	Drosophila melanog
19	16.4	78.1	2645	23	ABL13722	Drosophila melanog
20	16.4	78.1	3541	23	ABL20652	Drosophila melanog
21	16.4	78.1	6015	23	ABL12304	Drosophila melanog
22	16.4	78.1	6747	23	ABL11528	Drosophila melanog
23	16.2	77.1	359	22	ABA73205	Human foetal liver
24	16.2	77.1	359	22	ABA38638	Probe #17104 for g
25	16.2	77.1	359	22	AAK21640	Human brain expres
26	16.2	77.1	359	22	AAK47805	Human bone marrow
27	16.2	77.1	359	22	AAI26041	Probe #15974 for g
28	16.2	77.1	359	22	AAI53633	Probe #22319 used
29	16.2	77.1	359	24	ABS21812	Human genome-deriv
30	16.2	77.1	476	22	ABA43210	Human breast cell
31	16.2	77.1	476	22	ABA53638	Human foetal liver
32	16.2	77.1	476	22	ABA23390	Probe #1856 for ge
33	16.2	77.1	476	22	AAK01900	Human brain expres
34	16.2	77.1	476	22	AAK27357	Human bone marrow
35	16.2	77.1	476	22	AAI11930	Probe #1863 for ge
36	16.2	77.1	476	22	AAI32661	Probe #1947 used t
37	16.2	77.1	476	22	AAI01867	Probe #1858 used t
38	16.2	77.1	476	24	ABN01870	Human genome-deriv
39	16.2	77.1	476	24	ABN96380	Gene #2878 used to
40	16.2	77.1	597	22	ABA60667	Human foetal liver
41	16.2	77.1	597	22	ABA28761	Probe #7227 for ge
42	16.2	77.1	597	22	AAK08949	Human brain expres
43	16.2	77.1	597	22	AAK34840	Human bone marrow
44	16.2	77.1	597	22	AAI16846	Probe #6779 for ge
45	16.2	77.1	597	22	AAI40556	Probe #9242 used t

ALIGNMENTS

RESULT 1
AAZ36155
ID AAZ36155 standard; DNA; 21 BP.

AC AAZ36155;

DT 11-FEB-2000 (first entry)

DE PCR primer for DNA encoding cancer associated antigen KOC-1.

KW Cancer associated antigen; KOC-1; cancer; vaccine; CT7; PCR primer; sg.

OS Synthetic.

OS Homo sapiens.

PN WO9954738-A1.

PD 28-OCT-1999.

PF 16-MAR-1999; 99WO-US05766.

PR 17-APR-1998; 98US-0061709.

PA (LUDW-) LUDWIG INST CANCER RES.

PI Chen Y, Gure A, Tsang S, Stockert E, Jager E, Knuth A, Old LJ;

DR WPI; 2000-013284/01.

PT Nucleotides representing cancer-associated genes, used to develop

products for the diagnosis, monitoring and treatment of cancers

XX

Instant

CC can be used to monitor the progression of a cancer by carrying out the
CC detection at subsequent time points and comparing the results from the
CC different time points. CD4+ and/or CD8+ T-Cells isolated from a patient
CC are treated with P2, polynucleotides encoding P2 or antigen presenting
CC cells expressing P2 and then administered to the patient to inhibit
CC development of cancer.

XX Sequence 1740 BP; 526 A; 406 C; 417 G; 391 T; 0 other;

Query Match 100.0%; Score 21; DB 21; Length 1740;

Best Local Similarity 100.0%; Pred. No. 0.55; Indels 0; Gaps 0;

Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAAAGTATCTTCAGGAGGCC 21

DB 55 GAAAGTATCTTCAGGAGGCC 75

RESULT 3

ABL49254

ID ABL49254 standard; cDNA; 1740 BP.

XX ABL49254;

XX 01-MAY-2002 (first entry)

XX Human lung tumour L523S cDNA sequence SEQ ID NO:347.

XX Human; lung tumour; lung cancer; cytostatic; immunostimulant; vaccine;

XX immune response; ss.

XX Homo sapiens.

XX WO200200174-A2.

XX 03-JAN-2002.

XX 28-JUN-2001; 2001WO-US21065.

XX 28-JUN-2000; 2000US-0606421.

XX 02-AUG-2000; 2000US-0630940.

XX 21-AUG-2000; 2000US-0643597.

XX 15-SEP-2000; 2000US-0662786.

XX 09-OCT-2000; 2000US-0685696.

XX 12-DEC-2000; 2000US-0735705.

XX 07-MAY-2001; 2001US-0850716.

XX (CORI-) CORIXA CORP.

XX Wang T, Wang A, Skeiky YAW, Li SX, Kalos MD, Henderson RA;

XX McNeill PD, Fanger N, Retter MW, Warnerakis M, Fanger GR;

XX Vedvick TS, Carter D, Watanabe Y, Peckham DW;

XX WPI; 2002-090513/12.

XX P-PSDB; ABB74997.

XX Polynucleotides encoding lung tumor polypeptides, useful for treating

XX lung cancer or stimulating an immune response -

XX Example 2; Page 330; 374pp; English.

XX The present invention describes human lung tumour proteins. Human lung
XX tumour proteins and polynucleotides have cytostatic and immunostimulant
XX activities, and can be used in vaccine production. Compositions
XX comprising the lung tumour proteins, polynucleotides, antibodies,
XX fusion proteins, T cell populations, or antigen presenting cells that
XX express the lung tumour proteins are useful for treating lung cancer or
XX stimulating an immune response. ABL48959 to ABL49300 and ABB74946 to
XX ABB75070 represent sequences used in the exemplification of the present
XX invention.

XX Sequence 1740 BP; 526 A; 406 C; 417 G; 391 T; 0 other;

PS Claim 108; Page 13; 44pp; English.

XX PCR primers AA236155-56 were used to amplify a cancer associated antigen
CC gene designated KOC-1. The specification also describes a cancer
CC associated antigen designated CT7. The CT7 polynucleotide was isolated
CC from SK-MEL-37 melanoma cells. The polypeptide has some homology with
CC MAGE-10, limited to about 210 carboxy terminal amino acids. The amino
CC terminal of the protein has a repetitive pattern, with repeats rich in
CC serine, proline, glutamine and leucine, and an almost invariable core of
CC the peptide given in AAY43877. The CT7 polypeptide can be processed to
CC peptides which provoke lysis by cytolytic T cells. The polynucleotides
CC and polypeptides can be used for treating a cancerous condition and
CC screening for or diagnosing cancerous conditions. The cancer associated
CC antigens can be used as an immunogenic or vaccine composition with an
CC adjuvant, e.g. a cytokine, a saponin, or granulocyte macrophage-colony
CC stimulating factor (GM-CSF).

XX Sequence 21 BP; 7 A; 5 C; 5 G; 4 T; 0 other;

Query Match 100.0%; Score 21; DB 21; Length 21;

Best Local Similarity 100.0%; Pred. No. 0.26; Indels 0; Gaps 0;

Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAAAGTATCTTCAGGAGGCC 21

DB 1 GAAAGTATCTTCAGGAGGCC 21

RESULT 2

AAC66035

ID AAC66035 standard; cDNA; 1740 BP.

XX AAC66035;

XX 21-FEB-2001 (first entry)

XX Human lung cancer-associated cDNA antigen L523S.

XX Lung cancer; therapy; treatment; human; tumor; immunogenic; cytostatic;

XX vaccine; detection; ss.

XX Homo sapiens.

XX WO200061612-A2.

XX 19-OCT-2000.

XX 03-APR-2000; 2000WO-US08896.

XX 02-APR-1999; 93US-0285479.

XX 17-DEC-1999; 93US-0466396.

XX 30-DEC-1999; 93US-0476496.

XX 10-JAN-2000; 2000US-0480884.

XX 22-FEB-2000; 2000US-0510376.

XX (CORI-) CORIXA CORP.

XX Wang T, Fan L;

XX WPI; 2000-628399/60.

XX P-PSDB; AAB11365.

XX Isolated polypeptide comprising an immunogenic portion of a lung tumor

XX protein is used for detecting and monitoring progression of lung cancer

XX in a patient -

XX Claim 1a; Page 258-259; 261pp; English.

XX This invention describes a novel isolated polypeptide (I) which
XX comprising an immunogenic portion of a lung tumor protein or variant (P2)
XX which have cytostatic activity. The polypeptides and polynucleotides are
XX used in compositions and vaccines to inhibit the development of cancer,
XX especially lung cancer, in a patient. Methods described in the invention

Query Match 100.0%; Score 21; DB 24; Length 1740;
 Best Local Similarity 100.0%; Pred. No. 0.55;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAAAGTATCTTCAAGGACGCC 21
 |||||
 DB 55 GAAAGTATCTTCAAGGACGCC 75

RESULT 4

ABL49297
 ID ABL49297 standard; cDNA; 1743 BP.

XX ABL49297;

DT 01-MAY-2002 (first entry)

DE Human lung tumour L523S recombinant polynucleotide SEQ ID NO:447;

KW Human; lung tumour; lung cancer; cytostatic; immunostimulant; vaccine;
 immune response; ss.

OS Homo sapiens.

PN WO200200174-A2.

PD 03-JAN-2002.

PF 28-JUN-2001; 2001WO-US21065.

PR 28-JUN-2000; 2000US-0606421.

PR 02-AUG-2000; 2000US-0630940.

PR 21-AUG-2000; 2000US-0643597.

PR 15-SEP-2000; 2000US-0662786.

PR 09-OCT-2000; 2000US-0685696.

PR 12-DEC-2000; 2000US-0735705.

PR 07-MAY-2001; 2001US-0850716.

XX (CORI-) CORIXA CORP.

PI Wang T, Wang A, Skeiky YAW, Li SX, Kalos MD, Henderson RA;

PI McNeill PD, Fanger N, Retter MW, Marnerakis M, Fanger GR;

PI Vedvick TS, Carter D, Watanabe Y, Peckham DW;

XX P-PSDB; ABB75053.

DR WPI; 2002-090513/12.

DR P-PSDB; ABB75053.

PT Polynucleotides encoding lung tumor polypeptides, useful for treating
 lung cancer or stimulating an immune response -

PS Claim 1; Page 367; 374pp; English.

CC The present invention describes human lung tumour proteins. Human lung
 tumour proteins and polynucleotides have cytostatic and immunostimulant
 activities, and can be used in vaccine production. Compositions
 comprising the lung tumour proteins, polynucleotides, antibodies, that
 fusion proteins, T cell populations, or antigen presenting cells that
 express the lung tumour proteins are useful for treating lung cancer or
 stimulating an immune response. ABL48959 to ABL49300 and ABB74946 to
 ABB75070 represent sequences used in the exemplification of the present
 invention.

Sequence 1743 BP; 527 A; 406 C; 418 G; 392 T; 0 other;

Query Match 100.0%; Score 21; DB 24; Length 1743;
 Best Local Similarity 100.0%; Pred. No. 0.55;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAAAGTATCTTCAAGGACGCC 21

|||||
 DB 55 GAAAGTATCTTCAAGGACGCC 75

RESULT 5

ABL49299
 ID ABL49299 standard; cDNA; 1743 BP.

XX ABL49299;

DT 01-MAY-2002 (first entry)

DE Human lung tumour L523S recombinant polynucleotide SEQ ID NO:450.

KW Human; lung tumour; lung cancer; cytostatic; immunostimulant; vaccine;
 immune response; ss.

OS Homo sapiens.

PN WO200200174-A2.

PD 03-JAN-2002.

PF 28-JUN-2001; 2001WO-US21065.

PR 28-JUN-2000; 2000US-0606421.

PR 02-AUG-2000; 2000US-0630940.

PR 21-AUG-2000; 2000US-0643597.

PR 15-SEP-2000; 2000US-0662786.

PR 09-OCT-2000; 2000US-0685696.

PR 12-DEC-2000; 2000US-0735705.

PR 07-MAY-2001; 2001US-0850716.

XX (CORI-) CORIXA CORP.

PI Wang T, Wang A, Skeiky YAW, Li SX, Kalos MD, Henderson RA;

PI McNeill PD, Fanger N, Retter MW, Marnerakis M, Fanger GR;

PI Vedvick TS, Carter D, Watanabe Y, Peckham DW;

XX WPI; 2002-090513/12.

DR P-PSDB; ABB75054.

PT Polynucleotides encoding lung tumor polypeptides, useful for treating
 lung cancer or stimulating an immune response -

PS Claim 1; Page 370; 374pp; English.

CC The present invention describes human lung tumour proteins. Human lung
 tumour proteins and polynucleotides have cytostatic and immunostimulant
 activities, and can be used in vaccine production. Compositions
 comprising the lung tumour proteins, polynucleotides, antibodies,
 fusion proteins, T cell populations, or antigen presenting cells that
 express the lung tumour proteins are useful for treating lung cancer or
 stimulating an immune response. ABL48959 to ABL49300 and ABB74946 to
 ABB75070 represent sequences used in the exemplification of the present
 invention.

Sequence 1743 BP; 527 A; 406 C; 418 G; 392 T; 0 other;

Query Match 100.0%; Score 21; DB 24; Length 1743;
 Best Local Similarity 100.0%; Pred. No. 0.55;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAAAGTATCTTCAAGGACGCC 21

|||||
 DB 55 GAAAGTATCTTCAAGGACGCC 75

RESULT 6

ABL49283
 ID ABL49283 standard; cDNA; 1764 BP.

XX ABL49283;

DT 01-MAY-2002 (first entry)

DE Human lung tumour L523S recombinant polynucleotide SEQ ID NO:428.

us-09-270-437d-9.rng

Mon Apr 21 10:30:50 2003

XX Human; lung tumour; lung cancer; cytostatic; immunostimulant; vaccine;
 KW immune response; ss.
 XX Homo sapiens.
 OS WO200200174-A2.
 XX 03-JAN-2002.
 XX 28-JUN-2001; 2001WO-US21065.
 XX 28-JUN-2000; 2000US-0606421.
 PR 02-AUG-2000; 2000US-0630940.
 PR 21-AUG-2000; 2000US-0643597.
 PR 15-SEP-2000; 2000US-0662786.
 PR 09-OCT-2000; 2000US-0685696.
 PR 12-DEC-2000; 2000US-0735705.
 PR 07-MAY-2001; 2001US-0850716.
 XX (CORI-) CORIXA CORP.
 XX Wang T, Wang A, Skeiky YAW, Li SX, Kalos MD, Henderson RA;
 PI McNeill PD, Fanger N, Retter MW, Marnerakis M, Fanger GR;
 PI Vedvick TS, Carter D, Watanabe Y, Peckham DW;
 XX WPI; 2002-090513/12.
 DR P-PSDB; ABB75048.
 XX Polynucleotides encoding lung tumor polypeptides, useful for treating
 XX lung cancer or stimulating an immune response -
 PT Claim 1; Page 355-356; 374pp; English.
 XX The present invention describes human lung tumour proteins. Human lung
 CC tumour proteins and polynucleotides have cytostatic and immunostimulant
 CC activities, and can be used in vaccine production. Compositions
 CC comprising the lung tumour proteins, polynucleotides, antibodies,
 CC fusion proteins, T cell populations, or antigen presenting cells that
 CC express the lung tumour proteins are useful for treating lung cancer or
 CC stimulating an immune response. ABL48959 to ABL49300 and ABB74946 to
 CC ABB75070 represent sequences used in the exemplification of the present
 CC invention.
 XX Sequence 1764 BP; 534 A; 417 C; 419 G; 394 T; 0 other;
 SQ Query Match 100.0%; Score 21; DB 24; Length 1764;
 Best Local Similarity 100.0%; Pred. No. 0.55;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GAAAGTATCTTCAAGGACGCC 21
 DB 76 GAAAGTATCTTCAAGGACGCC 96
 RESULT 7
 AAZ36150
 ID AAZ36150 standard; DNA; 4159 BP.
 XX AAZ36150;
 AC AAZ36150;
 XX 11-FEB-2000 (first entry)
 DT DNA encoding cancer associated antigen KOC-1.
 DE Cancer associated antigen; KOC-1; cancer; vaccine; CT7; ss.
 XX Homo sapiens.
 OS WO9954738-A1.
 XX 28-OCT-1999.
 XX

PF 16-MAR-1999; 99WO-US05766.
 XX 17-APR-1998; 98US-0061709.
 XX (LUDW-) LUDWIG INST CANCER RES.
 XX Chen Y, Gure A, Tsang S, Stockert E, Jager E, Knuth A, Old LJ;
 PI WPI; 2000-013284/01.
 XX Nucleotides representing cancer-associated genes, used to develop
 PT products for the diagnosis, monitoring and treatment of cancers -
 PT Claim 88; Page 39-40; 44pp; English.
 XX The present sequence represents a cancer associated antigen gene
 CC designated KOC-1. The specification also describes a cancer associated
 CC antigen designated CT7. The CT7 polynucleotide was isolated from
 CC SK-MEL-37 melanoma cells. The polypeptide has some homology with
 CC MAGE-10, limited to about 210 carboxy terminal amino acids. The amino
 CC terminal of the protein has a repetitive pattern, with repeats rich in
 CC serine, proline, glutamine and leucine, and an almost invariable core of
 CC peptides which provoke lysis by cytolytic T cells. The polynucleotides
 CC and polypeptides can be used for treating a cancerous condition and
 CC screening for or diagnosing cancerous conditions. The cancer associated
 CC antigens can be used as an immunogenic or vaccine composition with an
 CC adjuvant, e.g. a cytokine, a saponin, or granulocyte macrophage-colony
 CC stimulating factor (GM-CSF).
 XX Sequence 4159 BP; 1281 A; 830 C; 851 G; 1181 T; 16 other;
 SQ Query Match 100.0%; Score 21; DB 21; Length 4159;
 Best Local Similarity 100.0%; Pred. No. 0.63;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GAAAGTATCTTCAAGGACGCC 21
 DB 305 GAAAGTATCTTCAAGGACGCC 325
 RESULT 8
 AAC65900
 ID AAC65900 standard; cDNA; 4181 BP.
 XX AAC65900;
 AC AAC65900;
 XX 21-FEB-2001 (first entry)
 DT Human lung cancer-associated cDNA L523S.
 DE Lung cancer; therapy; treatment; human; tumor; immunogenic; cytostatic;
 XX vaccine; detection; ss.
 KW Homo sapiens.
 OS WO2000061612-A2.
 XX 19-OCT-2000.
 XX 03-APR-2000; 2000WO-US08896.
 XX 02-APR-1999; 99US-0285479.
 PR 17-DEC-1999; 99US-0466396.
 PR 30-DEC-1999; 99US-0476496.
 PR 10-JAN-2000; 2000US-0480884.
 PR 22-FEB-2000; 2000US-0510376.
 XX (CORI-) CORIXA CORP.
 XX Wang T, Fan L;
 XX WPI; 2000-628399/60.
 DR

DR P-PSDB; AABL1328.
 XX Isolated polypeptide comprising an immunogenic portion of a lung tumor
 PT protein is used for detecting and monitoring progression of lung cancer
 PT in a patient -
 PS Claim 1a; Page 184-186; 261pp; English.
 XX
 CC This invention describes a novel isolated polypeptide (I) which
 CC comprising an immunogenic portion of a lung tumor protein or variant (P2)
 CC which have cytostatic activity. The polypeptides and polynucleotides are
 CC used in compositions and vaccines to inhibit the development of cancer,
 CC especially lung cancer, in a patient. Methods described in the invention
 CC can be used to monitor the progression of a cancer by carrying out the
 CC detection at subsequent time points and comparing the results from the
 CC different time points. CD4+ and/or CD8+ T-Cells isolated from a patient
 CC are treated with P2, polynucleotides encoding P2 or antigen presenting
 CC cells expressing P2 and then administered to the patient to inhibit
 CC development of cancer.
 XX
 SQ Sequence 4181 BP; 1303 A; 830 C; 851 G; 1181 T; 16 other;
 Query Match 100.0%; Score 21; DB 21; Length 4181;
 Best Local Similarity 100.0%; Pred. No. 0.63;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GAAAGTATCTTCAAGGAGCC 21
 Db 305 GAAAGTATCTTCAAGGAGCC 325
 RESULT 9
 ABL65405
 ID ABL65405 standard; DNA; 4181 BP.
 XX
 AC ABL65405;
 XX
 DT 15-MAY-2002 (first entry)
 XX
 DE Lung cancer related gene sequence SEQ ID NO:3742.
 XX
 KW Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;
 KW stomach; lung; prostate; pancreas; carcinoma; antitumor; cancerous;
 KW cytostatic; Gene therapy; antineoplastic; Wilms' tumour; adenocarcinoma;
 KW Gene; ds.
 OS
 OS Homo sapiens.
 XX
 PN WO200194629-A2.
 XX
 PD 13-DEC-2001.
 XX
 PF 30-MAY-2001; 2001WO-US10838.
 XX
 PR 05-JUN-2000; 2000US-209473P.
 PR 05-JUN-2000; 2000US-209531P.
 PR 18-SEP-2000; 2000US-233133P.
 PR 18-SEP-2000; 2000US-233617P.
 PR 20-SEP-2000; 2000US-234009P.
 PR 20-SEP-2000; 2000US-234034P.
 PR 22-SEP-2000; 2000US-234052P.
 PR 22-SEP-2000; 2000US-234509P.
 PR 22-SEP-2000; 2000US-234567P.
 PR 25-SEP-2000; 2000US-234923P.
 PR 25-SEP-2000; 2000US-234924P.
 PR 25-SEP-2000; 2000US-235077P.
 PR 25-SEP-2000; 2000US-235082P.
 PR 25-SEP-2000; 2000US-235134P.
 PR 26-SEP-2000; 2000US-235280P.
 PR 26-SEP-2000; 2000US-235637P.
 PR 26-SEP-2000; 2000US-235638P.
 PR 27-SEP-2000; 2000US-235711P.
 PR 27-SEP-2000; 2000US-235720P.

PR 27-SEP-2000; 2000US-235840P.
 PR 27-SEP-2000; 2000US-235863P.
 PR 28-SEP-2000; 2000US-236028P.
 PR 28-SEP-2000; 2000US-236032P.
 PR 28-SEP-2000; 2000US-236033P.
 PR 28-SEP-2000; 2000US-236034P.
 PR 28-SEP-2000; 2000US-236109P.
 PR 28-SEP-2000; 2000US-236111P.
 PR 29-SEP-2000; 2000US-236842P.
 PR 29-SEP-2000; 2000US-236891P.
 PR 02-OCT-2000; 2000US-237172P.
 PR 02-OCT-2000; 2000US-237173P.
 PR 02-OCT-2000; 2000US-237278P.
 PR 02-OCT-2000; 2000US-237294P.
 PR 02-OCT-2000; 2000US-237295P.
 PR 02-OCT-2000; 2000US-237316P.
 PR 03-OCT-2000; 2000US-237425P.
 PR 03-OCT-2000; 2000US-237598P.
 PR 03-OCT-2000; 2000US-237604P.
 PR 03-OCT-2000; 2000US-237606P.
 PR 03-OCT-2000; 2000US-237608P.
 PR 01-NOV-2000; 2000US-244867P.
 PR 01-NOV-2000; 2000US-245084P.
 XX
 PA (AVAL-) AVALON PHARM.
 XX
 XX Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;
 PI Soppet DR, Weaver Z;
 PI
 XX WPI; 2002-188264/24.
 XX
 XX Screening for anti-neoplastic agent involves exposing cells to a
 PT chemical agent to be tested for anti-neoplastic activity, and
 PT determining a change in expression of a gene of a signature gene set -
 XX
 PS Claim 1; SEQ ID 3742; 44pp; English.
 XX
 CC The present invention describes a method (M1) for screening for an
 CC anti-neoplastic agent. The method involves exposing cells to a chemical
 CC agent to be tested for anti-neoplastic activity, determining a change in
 CC expression of at least one gene (I) of a signature gene set, where (I)
 CC comprises a sequence (S) selected from 8447 sequences (given in ABL61664
 CC to ABL70110), or is at least 95% identical to (S), where a change in
 CC expression is indicative of anti-neoplastic activity. (I) has cytostatic
 CC activity and can be used in gene therapy. M1 can be used for screening
 CC an anti-neoplastic agent, and can be used for producing a product which
 CC is the data collected with respect to the anti-neoplastic agent as a
 CC result of M1, and the data is sufficient to convey the chemical
 CC structure and/or properties of the agent. M1 can be used in the
 CC treatment of cancer such as colon, breast, stomach, lung, thyroid,
 CC oesophageal, ovarian, kidney, prostate or pancreatic cancer,
 CC adenocarcinoma, carcinoma, clear cell cancer, infiltrating ductal cancer,
 CC infiltrating lobular cancer, squamous cell carcinoma, neuroendocrine
 CC carcinoma, papillary carcinoma and Wilms' tumour.
 XX
 SQ Sequence 4181 BP; 1303 A; 830 C; 851 G; 1181 T; 16 other;
 Query Match 100.0%; Score 21; DB 21; Length 4181;
 Best Local Similarity 100.0%; Pred. No. 0.63;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GAAAGTATCTTCAAGGAGCC 21
 Db 305 GAAAGTATCTTCAAGGAGCC 325
 RESULT 10
 ABL49119
 ID ABL49119 standard; cDNA; 4181 BP.
 XX
 XX ABL49119;
 AC
 XX
 DT 01-MAY-2002 (first entry)

us-09-270-437d-9.rng

Mon Apr 21 10:30:50 2003

XX DE Human lung tumour L523S cDNA sequence SEQ ID NO:175.
 XX KW Human; lung tumour; lung cancer; cytostatic; immunostimulant; vaccine;
 XX KW immune response; ss.
 XX OS Homo sapiens.
 XX PN WO200200174-A2.
 XX PD 03-JAN-2002.
 XX PF 28-JUN-2001; 2001WO-US21065.
 XX PR 28-JUN-2000; 2000US-0606421.
 XX PR 02-AUG-2000; 2000US-0630940.
 XX PR 21-AUG-2000; 2000US-0643597.
 XX PR 15-SEP-2000; 2000US-0662786.
 XX PR 09-OCT-2000; 2000US-0685696.
 XX PR 12-DEC-2000; 2000US-0735705.
 XX PR 07-MAY-2001; 2001US-0850716.
 XX PA (CORI-) CORIXA CORP.
 XX PI Wang T, Wang A, Skeiky YAW, Li SX, Kalos MD, Henderson RA;
 XX PI McNeill PD, Fanger N, Retter MW, Marnerakis M, Fanger GR;
 XX PI Vedvick TS, Carter D, Watanabe Y, Peckham DW;
 XX PI WPI; 2002-090513/12.
 XX DR P-PSDB; ABB74960.
 XX PT Polynucleotides encoding lung tumour polypeptides, useful for treating
 XX PT lung cancer or stimulating an immune response -
 XX PS Example 2; Page 266-267; 374pp; English.
 XX CC The present invention describes human lung tumour proteins. Human lung
 XX CC tumour proteins and polynucleotides have cytostatic and immunostimulant
 XX CC activities, and can be used in vaccine production. Compositions
 XX CC comprising the lung tumour proteins, polynucleotides, antibodies,
 XX CC fusion proteins, T cell populations, or antigen presenting cells that
 XX CC express the lung tumour proteins are useful for treating lung cancer or
 XX CC stimulating an immune response. ABL48959 to ABL49300 and ABB74946 to
 XX CC ABB75070 represent sequences used in the exemplification of the present
 XX CC invention.
 XX SQ Sequence 4181 BP; 1303 A; 830 C; 851 G; 1181 T; 16 other;
 Query Match 100.0%; Score 21; DB 24; Length 4181;
 Best Local Similarity 100.0%; Pred. No. 0.63;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GAAAGTATCTTCAAGGACGCC 21
 Db 305 GAAAGTATCTTCAAGGACGCC 325
 RESULT 11
 ABA99958
 ID ABA99958 standard; DNA; 4601 BP.
 XX AC ABA99958;
 XX DT 05-JUL-2002 (first entry)
 XX DE Human KOC DNA.
 XX KW KOC; promoter; human; cytostatic; dermatological; vulnary; tumour;
 XX KW chemotherapy; allogenic bone marrow transplant; pancreatic carcinoma;
 XX KW domain-containing protein overexpressed in cancer; immunostimulant;
 XX KW gene therapy; chronic pancreatitis; pluripotency; aging; wound healing;
 XX KW radiation therapy; skin regeneration; cell differentiation; cancer; ds.

OS Homo sapiens.
 XX PN WO200220036-A1.
 XX PD 14-MAR-2002.
 XX PF 01-AUG-2001; 2001WO-DE02948.
 XX PR 06-SEP-2000; 2000DE-1043964.
 XX PR 14-FEB-2001; 2001DE-1006829.
 XX PA (MUEL/) MUELLER F.
 XX PI Mueller F, Gress T, Adler G;
 XX PI WPI; 2002-292230/33.
 XX DR New promoter of the KOC gene, useful for diagnosis and treatment of
 XX PT cancer, for inducing cellular dedifferentiation and for identifying
 XX PT modulators -
 XX PS Claim 2; Fig 1a-b; 74pp; German.
 XX CC This invention describes a novel KOC (KH domain-containing protein
 XX CC overexpressed in cancer) promoter. The products of the invention have
 XX CC cytotatic, dermatological, vulnary and immunostimulant activity, are
 XX CC capable of modulating activity/expression of the KOC protein and can be
 XX CC used for gene therapy. The KOC promoter is also useful (i) for diagnosis
 XX CC of tumours that express KOC, particularly for differentiating between
 XX CC chronic pancreatitis and pancreatic carcinoma; (ii) to detect
 XX CC premalignant lesions of uncertain nature and to classify them for risk,
 XX CC and (iii) to evaluate treatments; (iv) to impart pluripotency to cells;
 XX CC (v) to prepare tissues or organs (by differentiation of stem cell
 XX CC populations); (vi) in high-dose chemotherapy; (vii) to improve ex vivo
 XX CC expansion of hematopoietic stem cells; (viii) to improve engraftment of
 XX CC allogenic bone marrow transplants; (ix) to slow down or reverse aging
 XX CC processes; (x) as prophylactic agents during chemotherapy/radiation
 XX CC therapy; (xi) to regenerate skin defects and accelerate wound healing,
 XX CC including for cosmetic purposes; and (xii) to immunise against malignant
 XX CC tumours or their precursor stages. The promoter can also be used (i) for
 XX CC isolation/enrichment/selective replication of stem cells; (ii) to
 XX CC determine the degree of (de)differentiation of cells and tissues; (iii)
 XX CC in gene therapy of cancer (for controlling protein expression); and (iv)
 XX CC to identify compounds, potentially useful in cancer treatment, that bind
 XX CC to the promoter and modulate its activity. The promoter makes possible
 XX CC generation of stem cells of any selected type. Since expression of KOC
 XX CC is not sensitive to other physiological states, e.g. inflammation, it
 XX CC represents a tumour marker with 100% sensitivity and specificity. This
 XX CC sequence represents the human KOC gene described in the disclosure of
 XX CC the invention.
 XX SQ Sequence 4601 BP; 1462 A; 906 C; 918 G; 1298 T; 17 other;
 Query Match 100.0%; Score 21; DB 24; Length 4601;
 Best Local Similarity 100.0%; Pred. No. 0.64;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GAAAGTATCTTCAAGGACGCC 21
 Db 725 GAAAGTATCTTCAAGGACGCC 745
 RESULT 12
 AAS86150
 ID AAS86150 standard; cDNA; 4264 BP.
 XX AC AAS86150;
 XX DT 13-FEB-2002 (first entry)
 XX DE DNA encoding novel human diagnostic protein #21954.
 XX KW Human; chromosome mapping; gene mapping; gene therapy; forensic;

KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
 OS Homo sapiens.
 XX WO200175067-A2.
 XX PD 11-OCT-2001.
 XX PF 30-MAR-2001; 2001WO-US08631.
 XX PR 31-MAR-2000; 2000US-0540217.
 XX PR 23-AUG-2000; 2000US-0649167.
 XX PA (HYSE-) HYSEQ INC.
 XX PI Drmanac RT, Liu C, Tang YT;
 XX DR WPI; 2001-639362/73.
 XX DR P-PSDB; ABG21963.
 XX PT New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity. -
 XX

PS Claim 1; SEQ ID No 21954; 103pp; English.
 XX
 CC The invention relates to isolated polynucleotide (I) and
 CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
 CC and gene mapping, and in recombinant production of (II). The
 CC polynucleotides are also used in diagnostics as expressed sequence tags
 CC for identifying expressed genes. (I) is useful in gene therapy techniques
 CC to restore normal activity of (II) or to treat disease states involving
 CC quantitating a polypeptide in tissue, as molecular weight markers and as
 CC imaging of sites expressing (II). (I) and (II) are useful for treating
 CC disorders involving aberrant protein expression or biological activity.
 CC The polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. AAS64197-AAS94564 represent novel human
 CC diagnostic coding sequences of the invention.
 CC Note: The sequence data for this patent did not appear in the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX

SQ Sequence 4264 BP; 1301 A; 854 C; 881 G; 1228 T; 0 other;
 Query Match 95.2%; Score 20; DB 23; Length 4264;
 Best Local Similarity 100.0%; Pred. No. 2.1;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 2 AAAGTATCTTCAGGACGCC 21
 |||||
 Db 398 AAAGTATCTTCAGGACGCC 417

RESULT 13
 AAS76779
 ID AAS76779 standard; cDNA; 1985 BP.
 XX
 AC AAS76779;
 XX
 DT 13-FEB-2002 (first entry)
 XX
 DE DNA encoding novel human diagnostic protein #12583.
 XX
 KW Human; chromosome mapping; gene mapping; gene therapy; forensics;
 KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
 XX

OS Homo sapiens.
 XX WO200175067-A2.
 XX PD 11-OCT-2001.
 XX PF 30-MAR-2001; 2001WO-US08631.
 XX PR 31-MAR-2000; 2000US-0540217.
 XX PR 23-AUG-2000; 2000US-0649167.
 XX PA (HYSE-) HYSEQ INC.
 XX PI Drmanac RT, Liu C, Tang YT;
 XX DR WPI; 2001-639362/73.
 XX DR P-PSDB; ABG21963.
 XX PT New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity. -
 XX

PS Claim 1; SEQ ID No 21954; 103pp; English.
 XX
 CC The invention relates to isolated polynucleotide (I) and
 CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
 CC and gene mapping, and in recombinant production of (II). The
 CC polynucleotides are also used in diagnostics as expressed sequence tags
 CC for identifying expressed genes. (I) is useful in gene therapy techniques
 CC to restore normal activity of (II) or to treat disease states involving
 CC quantitating a polypeptide in tissue, as molecular weight markers and as
 CC imaging of sites expressing (II). (I) and (II) are useful for treating
 CC disorders involving aberrant protein expression or biological activity.
 CC The polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. AAS64197-AAS94564 represent novel human
 CC diagnostic coding sequences of the invention.
 CC Note: The sequence data for this patent did not appear in the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX

SQ Sequence 4264 BP; 1301 A; 854 C; 881 G; 1228 T; 0 other;
 Query Match 95.2%; Score 20; DB 23; Length 4264;
 Best Local Similarity 100.0%; Pred. No. 2.1;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 2 AAAGTATCTTCAGGACGCC 21
 |||||
 Db 398 AAAGTATCTTCAGGACGCC 417

RESULT 13
 AAS76779
 ID AAS76779 standard; cDNA; 1985 BP.
 XX
 AC AAS76779;
 XX
 DT 13-FEB-2002 (first entry)
 XX
 DE DNA encoding novel human diagnostic protein #12583.
 XX
 KW Human; chromosome mapping; gene mapping; gene therapy; forensics;
 KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
 XX

OS Homo sapiens.
 XX WO200175067-A2.
 XX PD 11-OCT-2001.
 XX PF 30-MAR-2001; 2001WO-US08631.
 XX PR 31-MAR-2000; 2000US-0540217.
 XX PR 23-AUG-2000; 2000US-0649167.
 XX PA (HYSE-) HYSEQ INC.
 XX PI Drmanac RT, Liu C, Tang YT;
 XX DR WPI; 2001-639362/73.
 XX DR P-PSDB; ABG12592.
 XX PT New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity. -
 XX

PS Claim 1; SEQ ID No 12583; 103pp; English.
 XX
 CC The invention relates to isolated polynucleotide (I) and
 CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
 CC and gene mapping, and in recombinant production of (II). The
 CC polynucleotides are also used in diagnostics as expressed sequence tags
 CC for identifying expressed genes. (I) is useful in gene therapy techniques
 CC to restore normal activity of (II) or to treat disease states involving
 CC quantitating a polypeptide in tissue, as molecular weight markers and as
 CC imaging of sites expressing (II). (I) and (II) are useful for treating
 CC disorders involving aberrant protein expression or biological activity.
 CC The polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. AAS64197-AAS94564 represent novel human
 CC diagnostic coding sequences of the invention.
 CC Note: The sequence data for this patent did not appear in the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX

SQ Sequence 1985 BP; 611 A; 472 C; 460 G; 441 T; 1 other;
 Query Match 84.8%; Score 17.8; DB 23; Length 1985;
 Best Local Similarity 90.5%; Pred. No. 26;
 Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 1 GAAGTATCTTCAGGACGCC 21
 |||||
 Db 55 GAAGTATCTTCAGGACGCC 75

RESULT 14
 ABQ88221/c
 ID ABQ88221 standard; cDNA; 7661 BP.
 XX
 AC ABQ88221;
 XX
 DT 18-SEP-2002 (first entry)
 XX
 DE Human osteoblast differentiation related cDNA SEQ ID NO 128.
 XX
 KW Human; osteoblast; stem cell differentiation; bone tissue deposition;
 KW osteoporosis; osteopathic; ss.
 XX
 OS Homo sapiens.
 XX

OS Homo sapiens.
 XX WO200175067-A2.
 XX PD 11-OCT-2001.
 XX PF 30-MAR-2001; 2001WO-US08631.
 XX PR 31-MAR-2000; 2000US-0540217.
 XX PR 23-AUG-2000; 2000US-0649167.
 XX PA (HYSE-) HYSEQ INC.
 XX PI Drmanac RT, Liu C, Tang YT;
 XX DR WPI; 2001-639362/73.
 XX DR P-PSDB; ABG12592.
 XX PT New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity. -
 XX

PS Claim 1; SEQ ID No 12583; 103pp; English.
 XX
 CC The invention relates to isolated polynucleotide (I) and
 CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
 CC and gene mapping, and in recombinant production of (II). The
 CC polynucleotides are also used in diagnostics as expressed sequence tags
 CC for identifying expressed genes. (I) is useful in gene therapy techniques
 CC to restore normal activity of (II) or to treat disease states involving
 CC quantitating a polypeptide in tissue, as molecular weight markers and as
 CC imaging of sites expressing (II). (I) and (II) are useful for treating
 CC disorders involving aberrant protein expression or biological activity.
 CC The polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. AAS64197-AAS94564 represent novel human
 CC diagnostic coding sequences of the invention.
 CC Note: The sequence data for this patent did not appear in the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX

SQ Sequence 1985 BP; 611 A; 472 C; 460 G; 441 T; 1 other;
 Query Match 84.8%; Score 17.8; DB 23; Length 1985;
 Best Local Similarity 90.5%; Pred. No. 26;
 Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 1 GAAGTATCTTCAGGACGCC 21
 |||||
 Db 55 GAAGTATCTTCAGGACGCC 75

RESULT 14
 ABQ88221/c
 ID ABQ88221 standard; cDNA; 7661 BP.
 XX
 AC ABQ88221;
 XX
 DT 18-SEP-2002 (first entry)
 XX
 DE Human osteoblast differentiation related cDNA SEQ ID NO 128.
 XX
 KW Human; osteoblast; stem cell differentiation; bone tissue deposition;
 KW osteoporosis; osteopathic; ss.
 XX
 OS Homo sapiens.
 XX

OS Homo sapiens.
 XX WO200175067-A2.
 XX PD 11-OCT-2001.
 XX PF 30-MAR-2001; 2001WO-US08631.
 XX PR 31-MAR-2000; 2000US-0540217.
 XX PR 23-AUG-2000; 2000US-0649167.
 XX PA (HYSE-) HYSEQ INC.
 XX PI Drmanac RT, Liu C, Tang YT;
 XX DR WPI; 2001-639362/73.
 XX DR P-PSDB; ABG12592.
 XX PT New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity. -
 XX

PS Claim 1; SEQ ID No 12583; 103pp; English.
 XX
 CC The invention relates to isolated polynucleotide (I) and
 CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
 CC and gene mapping, and in recombinant production of (II). The
 CC polynucleotides are also used in diagnostics as expressed sequence tags
 CC for identifying expressed genes. (I) is useful in gene therapy techniques
 CC to restore normal activity of (II) or to treat disease states involving
 CC quantitating a polypeptide in tissue, as molecular weight markers and as
 CC imaging of sites expressing (II). (I) and (II) are useful for treating
 CC disorders involving aberrant protein expression or biological activity.
 CC The polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. AAS64197-AAS94564 represent novel human
 CC diagnostic coding sequences of the invention.
 CC Note: The sequence data for this patent did not appear in the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX

SQ Sequence 1985 BP; 611 A; 472 C; 460 G; 441 T; 1 other;
 Query Match 84.8%; Score 17.8; DB 23; Length 1985;
 Best Local Similarity 90.5%; Pred. No. 26;
 Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 1 GAAGTATCTTCAGGACGCC 21
 |||||
 Db 55 GAAGTATCTTCAGGACGCC 75

us-09-270-437d-9.rng

Mon Apr 21 10:30:50 2003

23-MAR-2001; 2001WO-US09231.
 23-MAR-2000; 2000US-191637P.
 11-JUL-2000; 2000US-0614150.
 (PEKE) PE CORP NY.
 Venter JC, Adams M, Li PWD, Myers EW;
 WPI; 2001-656860/75.
 P-PSDB; ABB69624.
 New isolated nucleic acid detection reagent for detecting 1000 or more genes from Drosophila and for elucidating cell signalling and cell-cell interactions -
 Claim 1; SEQ ID NO 35663; 21pp + Sequence Listing; English.
 The invention relates to an isolated nucleic acid detection reagent capable of detecting 1000 or more genes from Drosophila. The invention is useful in developmental biology and in elucidating cell signalling and cell-cell interactions in higher eukaryotes for the development of insecticides, therapeutics and pharmaceutical drugs. The invention discloses genomic DNA sequences (AB116176-AB130511), expressed DNA sequences (AB101840-AB116175) and the encoded proteins (AB157737-AB172072).
 The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.
 Sequence 557 BP; 135 A; 126 C; 139 G; 157 T; 0 other;
 Query Match 78.1%; Score 16.4; DB 23; Length 557;
 Best Local Similarity 94.4%; Pred. No. 1.1e+02;
 Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 4 AGTATCTTCAAGGAGCC 21
 DB 177 AGTATCTTCAAGGAGCC 160
 Search completed: April 18, 2003, 05:45:16
 Job time : 97.5455 secs

23-MAR-2001; 2001WO-US09231.
 23-MAR-2000; 2000US-191637P.
 11-JUL-2000; 2000US-0614150.
 (PEKE) PE CORP NY.
 Venter JC, Adams M, Li PWD, Myers EW;
 WPI; 2001-656860/75.
 P-PSDB; ABB69624.
 New isolated nucleic acid detection reagent for detecting 1000 or more genes from Drosophila and for elucidating cell signalling and cell-cell interactions -
 Claim 1; SEQ ID NO 35663; 21pp + Sequence Listing; English.
 The invention relates to an isolated nucleic acid detection reagent capable of detecting 1000 or more genes from Drosophila. The invention is useful in developmental biology and in elucidating cell signalling and cell-cell interactions in higher eukaryotes for the development of insecticides, therapeutics and pharmaceutical drugs. The invention discloses genomic DNA sequences (AB116176-AB130511), expressed DNA sequences (AB101840-AB116175) and the encoded proteins (AB157737-AB172072).
 The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.
 Sequence 557 BP; 135 A; 126 C; 139 G; 157 T; 0 other;
 Query Match 78.1%; Score 16.4; DB 23; Length 557;
 Best Local Similarity 94.4%; Pred. No. 1.1e+02;
 Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 4 AGTATCTTCAAGGAGCC 21
 DB 177 AGTATCTTCAAGGAGCC 160
 Search completed: April 18, 2003, 05:45:16
 Job time : 97.5455 secs

23-MAR-2001; 2001WO-US09231.
 23-MAR-2000; 2000US-191637P.
 11-JUL-2000; 2000US-0614150.
 (PEKE) PE CORP NY.
 Venter JC, Adams M, Li PWD, Myers EW;
 WPI; 2001-656860/75.
 P-PSDB; ABB69624.
 New isolated nucleic acid detection reagent for detecting 1000 or more genes from Drosophila and for elucidating cell signalling and cell-cell interactions -
 Claim 1; SEQ ID NO 35663; 21pp + Sequence Listing; English.
 The invention relates to an isolated nucleic acid detection reagent capable of detecting 1000 or more genes from Drosophila. The invention is useful in developmental biology and in elucidating cell signalling and cell-cell interactions in higher eukaryotes for the development of insecticides, therapeutics and pharmaceutical drugs. The invention discloses genomic DNA sequences (AB116176-AB130511), expressed DNA sequences (AB101840-AB116175) and the encoded proteins (AB157737-AB172072).
 The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.
 Sequence 557 BP; 135 A; 126 C; 139 G; 157 T; 0 other;
 Query Match 78.1%; Score 16.4; DB 23; Length 557;
 Best Local Similarity 94.4%; Pred. No. 1.1e+02;
 Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 4 AGTATCTTCAAGGAGCC 21
 DB 177 AGTATCTTCAAGGAGCC 160
 Search completed: April 18, 2003, 05:45:16
 Job time : 97.5455 secs

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 18, 2003, 04:53:41 ; Search time 20.5227 Seconds
(without alignments)
313.809 Million cell updates/sec

Title: US-09-270-437d-9

Perfect score: 21
Sequence: 1 gaaagtattcttcaaggagcgc 21

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 15333881 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA.*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq.*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq.*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq.*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq.*
5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq.*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	21	100.0	1740	4	US-09-643-597-347
2	21	100.0	4159	4	US-09-061-709-4
3	21	100.0	4181	4	US-09-643-597-175
4	15.8	75.2	936	2	US-08-288-663A-13
5	15.8	75.2	1194	2	US-08-288-663A-2
6	15.8	75.2	1228	2	US-08-288-663A-3
7	15.4	73.3	3207	1	US-08-045-806-1
8	15.4	73.3	3207	1	US-08-366-051B-1
9	15.4	73.3	5261	1	US-08-045-806-3
10	15.4	73.3	5261	1	US-08-366-051B-3
11	15.2	72.4	579	4	US-09-026-601-29
12	15.2	72.4	579	4	US-09-026-601-30
13	15.2	72.4	579	4	US-09-026-601-31
14	15.2	72.4	579	4	US-09-026-601-32
15	15.2	72.4	579	4	US-09-026-601-34
16	15.2	72.4	579	4	US-09-026-601-36
17	15.2	72.4	579	4	US-09-026-601-33
18	15.2	72.4	580	4	US-09-026-601-35
19	15.2	72.4	580	4	US-09-026-601-37
20	15.2	72.4	580	4	US-09-026-601-38
21	15.2	72.4	4422	4	US-09-166-350-1
22	15.2	72.4	7207	4	US-09-462-606-3
23	14.8	70.5	580	4	US-09-026-601-40
24	14.8	70.5	642	4	US-08-998-416-1126
25	14.6	69.5	399	4	US-09-134-001C-1761
26	14.6	69.5	433	1	US-08-664-596B-1
27	14.6	69.5	433	1	US-08-738-367-1

C 28	14.6	69.5	733	1	US-08-738-367-6	Sequence 6, Appli
C 29	14.6	69.5	1083	4	US-09-276-531-84	Sequence 84, Appl
C 30	14.6	69.5	1498	2	US-08-818-024-2	Sequence 2, Appli
C 31	14.6	69.5	1498	4	US-09-334-775A-2	Sequence 2, Appli
C 32	14.6	69.5	1654	4	US-09-347-798-9	Sequence 9, Appli
C 33	14.6	69.5	2175	2	US-08-389-564B-4	Sequence 4, Appli
C 34	14.6	69.5	2175	3	US-08-466-047B-4	Sequence 4, Appli
C 35	14.6	69.5	8224	2	US-09-010-398-14	Sequence 14, Appl
C 36	14.6	69.5	8224	4	US-09-366-260-14	Sequence 14, Appl
C 37	14.6	69.5	11663	1	US-08-446-932-1	Sequence 1, Appli
C 38	14.6	69.5	11663	1	US-08-801-263A-1	Sequence 1, Appli
C 39	14.6	69.5	11663	1	US-08-801-263A-7	Sequence 7, Appli
C 40	14.6	69.5	11663	3	US-09-102-248-1	Sequence 1, Appli
C 41	14.6	69.5	11663	3	US-09-102-248-7	Sequence 7, Appli
C 42	14.4	68.6	587	4	US-09-026-601-39	Sequence 39, Appl
C 43	14.4	68.6	662	4	US-09-221-017B-454	Sequence 454, App
C 44	14.4	68.6	2100	4	US-09-836-992-2	Sequence 2, Appli
C 45	14.4	68.6	2877	4	US-09-619-353-1	Sequence 1, Appli

ALIGNMENTS

RESULT 1
US-09-643-597-347
; Sequence 347, Application US/09643597
; Patent No. 6426072
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Fan, Liqun
; APPLICANT: Kalos, Michael D.
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Hosken, Nancy
; APPLICANT: Fanger, Gary R.
; APPLICANT: Li, Samuel X.
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Henderson, Robert A.
; APPLICANT: McNeill, Patricia D.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.455C11
; CURRENT APPLICATION NUMBER: US/09/643,597
; CURRENT FILING DATE: 2000-08-21
; NUMBER OF SEQ ID NOS: 369
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 347
; LENGTH: 1740
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-643-597-347

Query Match 100.0%; Score 21; DB 4; Length 1740;
Best Local Similarity 100.0%; Pred. No. 0.061;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GAAAGTATCTTCAAGGAGGCC 21
|||||
Db 55 GAAAGTATCTTCAAGGAGGCC 75

RESULT 2
US-09-061-709-4
; Sequence 4, Application US/09061709B
; Patent No. 6297364
; GENERAL INFORMATION:
; APPLICANT: Chen, Yao-Tseng
; APPLICANT: Gure, Ali
; APPLICANT: Tsang, Solam
; APPLICANT: Stockert, Elisabeth
; APPLICANT: Jager, Elke
; APPLICANT: Knuth, Alexander
; APPLICANT: Old, Lloyd J.

us-09-270-437d-9.rni

Mon Apr 21 10:30:52 2003

TITLE OF INVENTION: Isolated Nucleic Acid Molecules Encoding Cancer Associated Antigen, The Antigens Per Se, And Uses Thereof

FILE REFERENCE: LUD 5538

CURRENT APPLICATION NUMBER: US/09/061,709B

CURRENT FILING DATE: 1998-04-17

NUMBER OF SEQ ID NOS: 8

SEQ ID NO 4

LENGTH: 4159

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:

US-09-061-709-4

Query Match 100.0%; Score 21; DB 4; Length 4159;

Best Local Similarity 100.0%; Pred. No. 0.072;

Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAAAGTATCTTCAAGGAGGCC 21

Db 305 GAAAGTATCTTCAAGGAGGCC 325

RESULT 3

US-09-643-597-175

Sequence 175, Application US/09643597

Patent No. 6426072

GENERAL INFORMATION:

APPLICANT: Wang, Tongtong

APPLICANT: Fan, Liqun

APPLICANT: Kalos, Michael D.

APPLICANT: Bangur, Chaitanya S.

APPLICANT: Hosken, Nancy

APPLICANT: Fanger, Gary R.

APPLICANT: Li, Samuel X.

APPLICANT: Wang, Aijun

APPLICANT: Skeiky, Yasir A.W.

APPLICANT: Henderson, Robert A.

APPLICANT: McNeill, Patricia D.

TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY

TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER

FILE REFERENCE: 210121.455C11

CURRENT APPLICATION NUMBER: US/09/643,597

CURRENT FILING DATE: 2000-08-21

NUMBER OF SEQ ID NOS: 369

SOFTWARE: FastSeq for Windows Version 3.0

SEQ ID NO 175

LENGTH: 4181

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: unsure

LOCATION: (3347)

OTHER INFORMATION: n=A,T,C or G

NAME/KEY: unsure

LOCATION: (3502)

OTHER INFORMATION: n=A,T,C or G

NAME/KEY: unsure

LOCATION: (3506)

OTHER INFORMATION: n=A,T,C or G

NAME/KEY: unsure

LOCATION: (3520)

OTHER INFORMATION: n=A,T,C or G

NAME/KEY: unsure

LOCATION: (3538)

OTHER INFORMATION: n=A,T,C or G

NAME/KEY: unsure

LOCATION: (3549)

OTHER INFORMATION: n=A,T,C or G

NAME/KEY: unsure

LOCATION: (3646)

OTHER INFORMATION: n=A,T,C or G

NAME/KEY: unsure

LOCATION: (3940)

OTHER INFORMATION: n=A,T,C or G

NAME/KEY: unsure

LOCATION: (3968)

OTHER INFORMATION: n=A,T,C or G

NAME/KEY: unsure

LOCATION: (3974)

OTHER INFORMATION: n=A,T,C or G

NAME/KEY: unsure

LOCATION: (4036)

OTHER INFORMATION: n=A,T,C or G

NAME/KEY: unsure

LOCATION: (4056)

OTHER INFORMATION: n=A,T,C or G

NAME/KEY: unsure

LOCATION: (4062)

OTHER INFORMATION: n=A,T,C or G

NAME/KEY: unsure

LOCATION: (4080)

OTHER INFORMATION: n=A,T,C or G

NAME/KEY: unsure

LOCATION: (4088)

OTHER INFORMATION: n=A,T,C or G

NAME/KEY: unsure

LOCATION: (4115)

OTHER INFORMATION: n=A,T,C or G

US-09-643-597-175

Query Match 100.0%; Score 21; DB 4; Length 4181;

Best Local Similarity 100.0%; Pred. No. 0.072;

Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAAAGTATCTTCAAGGAGGCC 21

Db 305 GAAAGTATCTTCAAGGAGGCC 325

RESULT 4

US-08-288-663A-13

Sequence 13, Application US/08288663A

Patent No. 5879896

GENERAL INFORMATION:

APPLICANT: HINUMA, Shuji

APPLICANT: HOSoya, Masaki

APPLICANT: ONDA, Haruo

TITLE OF INVENTION: HUMAN TRH RECEPTOR, ITS PRODUCTION

TITLE OF INVENTION: AND USE

NUMBER OF SEQUENCES: 15

CORRESPONDENCE ADDRESS:

ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSMAN, LLP

STREET: 130 Water Street

CITY: Boston

STATE: MA

COUNTRY: USA

ZIP: 02109

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: FastSeq for Windows Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/288,663A

FILING DATE: 09-AUG-1994

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 198309/1993

FILING DATE: 10-AUG-1993

APPLICATION NUMBER: 286986/1993

FILING DATE: 16-NOV-1993

APPLICATION NUMBER: 325215/1993

FILING DATE: 22-DEC-1993

ATTORNEY/AGENT INFORMATION:

NAME: Resnick, David S

REGISTRATION NUMBER: 34,235

REFERENCE/DOCKET NUMBER: 44612
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-523-3400
TELEFAX: 617-523-6440
TELEX:
INFORMATION FOR SEQ ID NO: 13:
SEQUENCE CHARACTERISTICS:
LENGTH: 936 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-288-663A-13

Query Match 75.2%; Score 15.8; DB 2; Length 936;
Best Local Similarity 89.5%; Pred. No. 29;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 AAAGTATCTTCAAGGACG 20
Db 781 ACAGTATCTTCAAGGAGC 799

RESULT 5
US-08-288-663A-2
Sequence 2, Application US/08288663A
Patent No. 5879896
GENERAL INFORMATION:
APPLICANT: HINUMA, Shuji
APPLICANT: HOSOYA, Masaki
TITLE OF INVENTION: HUMAN TRH RECEPTOR, ITS PRODUCTION
TITLE OF INVENTION: AND USE
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSMAN, LLP
STREET: 130 Water Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/288,663A
FILING DATE: 09-AUG-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 198309/1993
FILING DATE: 10-AUG-1993
APPLICATION NUMBER: 286986/1993
FILING DATE: 16-NOV-1993
APPLICATION NUMBER: 325215/1993
FILING DATE: 22-DEC-1993
ATTORNEY/AGENT INFORMATION:
NAME: Resnick, David S
REGISTRATION NUMBER: 34,235
REFERENCE/DOCKET NUMBER: 44612
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-523-3400
TELEFAX: 617-523-6440
TELEX:
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1194 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-288-663A-2

Query Match 75.2%; Score 15.8; DB 2; Length 1194;
Best Local Similarity 89.5%; Pred. No. 30;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy 2 AAAGTATCTTCAAGGACG 20
Db 769 ACAGTATCTTCAAGGAGC 787

RESULT 6
US-08-288-663A-3
Sequence 3, Application US/08288663A
Patent No. 5879896
GENERAL INFORMATION:
APPLICANT: HINUMA, Shuji
APPLICANT: HOSOYA, Masaki
TITLE OF INVENTION: HUMAN TRH RECEPTOR, ITS PRODUCTION
TITLE OF INVENTION: AND USE
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSMAN, LLP
STREET: 130 Water Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/288,663A
FILING DATE: 09-AUG-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 198309/1993
FILING DATE: 10-AUG-1993
APPLICATION NUMBER: 286986/1993
FILING DATE: 16-NOV-1993
APPLICATION NUMBER: 325215/1993
FILING DATE: 22-DEC-1993
ATTORNEY/AGENT INFORMATION:
NAME: Resnick, David S
REGISTRATION NUMBER: 34,235
REFERENCE/DOCKET NUMBER: 44612
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-523-3400
TELEFAX: 617-523-6440
TELEX:
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1228 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-288-663A-3

Query Match 75.2%; Score 15.8; DB 2; Length 1228;
Best Local Similarity 89.5%; Pred. No. 30;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 AAAGTATCTTCAAGGACG 20
Db 781 ACAGTATCTTCAAGGAGC 799

RESULT 7
US-08-045-806-1
Sequence 1, Application US/08045806
Patent No. 5378822

Mon Apr 21 10:30:52 2003

```

;
; GENERAL INFORMATION:
; APPLICANT: Bradfield, Christopher Alan
; APPLICANT: Dolwick, Kristin Marie
; APPLICANT: Poland, Alan
; TITLE OF INVENTION: Ah Receptor cDNA and Method of
; TITLE OF INVENTION: Determining Human Risks To Environmental Pollutants
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Tilton, Fallon, Lungmus & Chestnut
; STREET: 100 South Wacker Drive, Suite 960
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606-4002
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/045,806
; FILING DATE: 19930408
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Pentress, Susan B.
; REGISTRATION NUMBER: 31,327
; REFERENCE/DOCKET NUMBER: NU-9207
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (312)-456-8000
; TELEFAX: (312)-456-7776
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3207 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: double
; TOPOLOGY: unknown
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..2415
; US-08-045-806-1
;
; Query Match 73.3%; Score 15.4; DB 1; Length 3207;
; Best Local Similarity 94.1%; Pred. No. 60;
; Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
;
QY 2 AAAGTATCTTCAAGGAC 18
|||
Db 711 AAAGTATCTTCATGGAC 727

RESULT 8
US-08-366-051B-1
; Sequence 1, Application US/08366051B
; Patent No. 5650283
; GENERAL INFORMATION:
; APPLICANT: Bradfield, Christopher A.
; APPLICANT: Dolwick, Kristin M.
; APPLICANT: Carver, Lucy A.
; TITLE OF INVENTION: Ah Receptor cDNAs and Genetically
; TITLE OF INVENTION: Engineered Cells for Detecting Agonists to the Ah
; TITLE OF INVENTION: Receptor
; NUMBER OF SEQUENCES: 36
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Tilton, Fallon, Lungmus & Chestnut
; STREET: 100 South Wacker Drive, Suite 960
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606-4002
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible

;
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/045,806
; FILING DATE: 19930408
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Pentress, Susan B.
; REGISTRATION NUMBER: 31,327
; REFERENCE/DOCKET NUMBER: NU-9207
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (312)-456-8000
; TELEFAX: (312)-456-7776
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3207 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: double
; TOPOLOGY: unknown
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..2415
; US-08-045-806-1
;
; Query Match 73.3%; Score 15.4; DB 1; Length 3207;
; Best Local Similarity 94.1%; Pred. No. 60;
; Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
;
QY 2 AAAGTATCTTCAAGGAC 18
|||
Db 711 AAAGTATCTTCATGGAC 727

RESULT 9
US-08-045-806-3
; Sequence 3, Application US/08045806
; Patent No. 5378822
; GENERAL INFORMATION:
; APPLICANT: Bradfield, Christopher Alan
; APPLICANT: Dolwick, Kristin Marie
; APPLICANT: Poland, Alan
; TITLE OF INVENTION: Ah Receptor cDNA and Method of
; TITLE OF INVENTION: Determining Human Risks To Environmental Pollutants
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Tilton, Fallon, Lungmus & Chestnut
; STREET: 100 South Wacker Drive, Suite 960
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606-4002
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/045,806
; FILING DATE: 19930408
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Pentress, Susan B.
; REGISTRATION NUMBER: 31,327
; REFERENCE/DOCKET NUMBER: NU-9207
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (312)-456-8000
; TELEFAX: (312)-456-7776
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5261 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: double
;
; Query Match 73.3%; Score 15.4; DB 1; Length 3207;
; Best Local Similarity 94.1%; Pred. No. 60;
; Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
;
QY 2 AAAGTATCTTCAAGGAC 18
|||
Db 711 AAAGTATCTTCATGGAC 727
```

;; TOPOLOGY: unknown
;; MOLECULE TYPE: cDNA
;; FEATURE:
;; NAME/KEY: CDS
;; LOCATION: 383..2927
US-08-045-806-3

Query Match 73.3%; Score 15.4; DB 1; Length 5261;
Best Local Similarity 94.1%; Pred. No. 66;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 AAAGTATCTTCAAGGAC 18
|||||
Db 1111 AAAGTATCTTCAAGGAC 1127

RESULT 10

US-08-366-051B-3
; Sequence 3, Application US/08366051B
; Patent No. 5650283
; GENERAL INFORMATION:
; APPLICANT: Bradfield, Christopher A.
; APPLICANT: Dolwick, Kristin M.
; APPLICANT: Carver, Lucy A.
; TITLE OF INVENTION: Ah Receptor cDNAs and Genetically
; TITLE OF INVENTION: Engineered Cells for Detecting Agonists to the Ah
; TITLE OF INVENTION: Receptor
; NUMBER OF SEQUENCES: 36
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Tilton, Fallon, Lungmus & Chestnut
; STREET: 100 South Wacker Drive, Suite 960
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606-4002
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/366,051B
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Tilton, Timothy L.
; REGISTRATION NUMBER: 16,926
; REFERENCE/DOCKET NUMBER: NU-9207-CIP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (312)-456-8000
; TELEFAX: (312)-456-7776
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5261 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: unknown
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 383..2927
US-08-366-051B-3

Query Match 73.3%; Score 15.4; DB 1; Length 5261;
Best Local Similarity 94.1%; Pred. No. 66;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 AAAGTATCTTCAAGGAC 18
|||||
Db 1111 AAAGTATCTTCAAGGAC 1127

RESULT 11

US-09-026-601-29/c
; Sequence 29, Application US/09026601
; Patent No. 6358680
; GENERAL INFORMATION:
; APPLICANT: Beck, James J.
; TITLE OF INVENTION: Detection of Wheat and Barley Fungal
; TITLE OF INVENTION: Pathogens Using the Polymerase Chain Reaction
; NUMBER OF SEQUENCES: 41
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 6358680artis Corporation
; STREET: 3054 Cornwallis Road
; CITY: Research Triangle Park
; STATE: No. 6358680th Carolina
; COUNTRY: USA
; ZIP: 27709
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/026,601
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Meigs, J. Timothy
; REGISTRATION NUMBER: 38,241
; REFERENCE/DOCKET NUMBER: CGC 1984
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 919-541-8587
; TELEFAX: 919-541-8689
; INFORMATION FOR SEQ ID NO: 29:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 579 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; ORIGINAL SOURCE:
; ORGANISM: Pyrenophora tritici-repentis
; INDIVIDUAL ISOLATE: consensus sequence
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 1..30
; OTHER INFORMATION: /note= "3' end of small subunit
; OTHER INFORMATION: rRNA gene"
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 31..208
; OTHER INFORMATION: /note= "ITS1"
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 209..365
; OTHER INFORMATION: /note= "5.8S rRNA gene"
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 366..526
; OTHER INFORMATION: /note= "ITS2"
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 527..579
; OTHER INFORMATION: /note= "5' end of large subunit
; OTHER INFORMATION: rRNA gene"
US-09-026-601-29

Query Match 72.4%; Score 15.2; DB 4; Length 579;
Best Local Similarity 85.0%; Pred. No. 54;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 GAAAGTATCTTCAAGGACG 20
|||||
Db 515 GAAAGTATCTTCAAGGACG 496

us-09-270-437d-9.rni

Mon Apr 21 10:30:52 2003

QY 1 GAAAGTATCTTCAAGGACGC 20
 Db 515 GAAAGAAGCTTCATGGACGC 496
 RESULT 13
 US-09-026-601-31/c
 ; Sequence 31, Application US/09026601
 ; Patent No. 6358680
 ; GENERAL INFORMATION:
 ; APPLICANT: Beck, James J.
 ; TITLE OF INVENTION: Detection of wheat and Barley Fungal
 ; TITLE OF INVENTION: Pathogens Using the Polymerase Chain Reaction
 ; NUMBER OF SEQUENCES: 41
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: No. 6358680artis Corporation
 ; STREET: 3054 Cornwallis Road
 ; CITY: Research Triangle Park
 ; STATE: No. 6358680th Carolina
 ; COUNTRY: USA
 ; ZIP: 27709
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent in Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/026,601
 ; FILING DATE:
 ; CLASSIFICATION:
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Meigs, J. Timothy
 ; REGISTRATION NUMBER: 38,241
 ; REFERENCE/DOCKET NUMBER: CGC 1984
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 919-541-8587
 ; TELEFAX: 919-541-8689
 ; INFORMATION FOR SEQ ID NO: 31:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 579 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: DNA (genomic)
 ; ORIGINAL SOURCE:
 ; ORGANISM: Pyrenophora tritici-repentis
 ; INDIVIDUAL ISOLATE: 119-2-3
 ; IMMEDIATE SOURCE:
 ; CLONE: 2-2
 ; FEATURE:
 ; NAME/KEY: misc feature
 ; LOCATION: 1..30
 ; OTHER INFORMATION: /note= "3' end of small subunit
 ; OTHER INFORMATION: rRNA gene"
 ; FEATURE:
 ; NAME/KEY: misc feature
 ; LOCATION: 31..208
 ; OTHER INFORMATION: /note= "ITS1"
 ; FEATURE:
 ; NAME/KEY: misc feature
 ; LOCATION: 209..365
 ; OTHER INFORMATION: /note= "5.8S rRNA gene"
 ; FEATURE:
 ; NAME/KEY: misc feature
 ; LOCATION: 366..526
 ; OTHER INFORMATION: /note= "ITS2"
 ; FEATURE:
 ; NAME/KEY: misc feature
 ; LOCATION: 527..579
 ; OTHER INFORMATION: /note= "5' end of large subunit
 ; OTHER INFORMATION: rRNA gene"
 ; US-09-026-601-31

RESULT 12
 US-09-026-601-30/c
 ; Sequence 30, Application US/09026601
 ; Patent No. 6358680
 ; GENERAL INFORMATION:
 ; APPLICANT: Beck, James J.
 ; TITLE OF INVENTION: Detection of wheat and Barley Fungal
 ; TITLE OF INVENTION: Pathogens Using the Polymerase Chain Reaction
 ; NUMBER OF SEQUENCES: 41
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: No. 6358680artis Corporation
 ; STREET: 3054 Cornwallis Road
 ; CITY: Research Triangle Park
 ; STATE: No. 6358680th Carolina
 ; COUNTRY: USA
 ; ZIP: 27709
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent in Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/026,601
 ; FILING DATE:
 ; CLASSIFICATION:
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Meigs, J. Timothy
 ; REGISTRATION NUMBER: 38,241
 ; REFERENCE/DOCKET NUMBER: CGC 1984
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 919-541-8587
 ; TELEFAX: 919-541-8689
 ; INFORMATION FOR SEQ ID NO: 30:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 579 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: DNA (genomic)
 ; ORIGINAL SOURCE:
 ; ORGANISM: Pyrenophora tritici-repentis
 ; INDIVIDUAL ISOLATE: 6715
 ; IMMEDIATE SOURCE:
 ; CLONE: 2 and 4 (consensus)
 ; FEATURE:
 ; NAME/KEY: misc feature
 ; LOCATION: 1..30
 ; OTHER INFORMATION: /note= "3' end of small subunit
 ; OTHER INFORMATION: rRNA gene"
 ; FEATURE:
 ; NAME/KEY: misc feature
 ; LOCATION: 31..208
 ; OTHER INFORMATION: /note= "ITS1"
 ; FEATURE:
 ; NAME/KEY: misc feature
 ; LOCATION: 209..365
 ; OTHER INFORMATION: /note= "5.8S rRNA gene"
 ; FEATURE:
 ; NAME/KEY: misc feature
 ; LOCATION: 366..526
 ; OTHER INFORMATION: /note= "ITS2"
 ; FEATURE:
 ; NAME/KEY: misc feature
 ; LOCATION: 527..579
 ; OTHER INFORMATION: /note= "5' end of large subunit
 ; OTHER INFORMATION: rRNA gene"
 ; US-09-026-601-30
 Query Match 72.4%; Score 15.2; DB 4; Length 579;
 Best Local Similarity 85.0%; Pred. No. 54;
 Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```
Query Match          72.4%; Score 15.2; DB 4; Length 579;
Best Local Similarity 85.0%; Pred. No. 54;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 GAAAGTATCTTCAAGGACGC 20
    ||||| ||||| |||||
Db 515 GAAAGAAGCTTCATGACGC 496

RESULT 14
US-09-026-601-32/c
; Sequence 32, Application US/09026601
; Patent No. 6358680
; GENERAL INFORMATION:
; APPLICANT: Beck, James J.
; TITLE OF INVENTION: Detection of Wheat and Barley Fungal
; NUMBER OF SEQUENCES: 41
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 6358680artis Corporation
; STREET: 3054 Cornwallis Road
; CITY: Research Triangle Park
; STATE: No. 6358680th Carolina
; COUNTRY: USA
; ZIP: 27709
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; FILING DATE:
; APPLICATION NUMBER: US/09/026.601
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Meigs, J. Timothy
; REGISTRATION NUMBER: 38,241
; REFERENCE/DOCKET NUMBER: CGC 1984
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 919-541-8587
; TELEFAX: 919-541-8689
; INFORMATION FOR SEQ ID NO: 32:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 579 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; ORIGINAL SOURCE:
; ORGANISM: Pyrenophora tritici-repentis
; INDIVIDUAL ISOLATE: DL22
; IMMEDIATE SOURCE:
; CLONE: 1-1
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1..30
; OTHER INFORMATION: /note= "3' end of small subunit"
; OTHER INFORMATION: rRNA gene"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 31..208
; OTHER INFORMATION: /note= "ITS1"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 209..365
; OTHER INFORMATION: /note= "5.8S rRNA gene"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 366..526
; OTHER INFORMATION: /note= "ITS2"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 527..579

Query Match          72.4%; Score 15.2; DB 4; Length 579;
Best Local Similarity 85.0%; Pred. No. 54;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 GAAAGTATCTTCAAGGACGC 20
    ||||| ||||| |||||
Db 515 GAAAGAAGCTTCATGACGC 496

RESULT 15
US-09-026-601-34/c
; Sequence 34, Application US/09026601
; Patent No. 6358680
; GENERAL INFORMATION:
; APPLICANT: Beck, James J.
; TITLE OF INVENTION: Detection of Wheat and Barley Fungal
; NUMBER OF SEQUENCES: 41
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 6358680artis Corporation
; STREET: 3054 Cornwallis Road
; CITY: Research Triangle Park
; STATE: No. 6358680th Carolina
; COUNTRY: USA
; ZIP: 27709
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; FILING DATE:
; APPLICATION NUMBER: US/09/026.601
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Meigs, J. Timothy
; REGISTRATION NUMBER: 38,241
; REFERENCE/DOCKET NUMBER: CGC 1984
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 919-541-8587
; TELEFAX: 919-541-8689
; INFORMATION FOR SEQ ID NO: 34:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 579 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; ORIGINAL SOURCE:
; ORGANISM: Pyrenophora tritici-repentis
; INDIVIDUAL ISOLATE: 44184
; IMMEDIATE SOURCE:
; CLONE: 3-1
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1..30
; OTHER INFORMATION: /note= "3' end of small subunit"
; OTHER INFORMATION: rRNA gene"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 31..208
; OTHER INFORMATION: /note= "ITS1"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 209..365
; OTHER INFORMATION: /note= "5.8S rRNA gene"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 366..526
```

us-09-270-437d-9.rni

Mon Apr 21 10:30:52 2003

```

; OTHER INFORMATION: /note= "ITS2"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 527..579
; OTHER INFORMATION: /note= "5' end of large subunit
; OTHER INFORMATION: rRNA gene"
; US-09-026-601-34
Query Match      72.4%; Score 15.2; DB 4; Length 579;
Best Local Similarity 85.0%; Pred. No. 54;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 GAAAGTATCTTCAAGGACGC 20
   ||||| ||||| ||||| |||||
Db 515 GAAAGAGCTTCATGGACGC 496

Search completed: April 18, 2003, 07:34:12
Job time : 25.5227 secs

```

GenCore version 5.1.4 p5 4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 19, 2003, 06:35:31 ; Search time 87.1818 Seconds
(without alignment)
241.975 Million cell updates/sec

Title: US-09-270-437D-9

Perfect score: 21
Sequence: 1 gaaagtattctcaaggagcc 21

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 639749 seqs, 502280978 residues

Total number of hits satisfying chosen parameters: 1279498

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA.*

- 1: /cgn2_6/ptodata/1/pubna/US07_PUBCOMB.seq.*
- 2: /cgn2_6/ptodata/1/pubna/PCT_NEW_PUB.seq.*
- 3: /cgn2_6/ptodata/1/pubna/US06_NEW_PUB.seq.*
- 4: /cgn2_6/ptodata/1/pubna/US06_PUBCOMB.seq.*
- 5: /cgn2_6/ptodata/1/pubna/US07_NEW_PUB.seq.*
- 6: /cgn2_6/ptodata/1/pubna/PCTUS_PUBCOMB.seq.*
- 7: /cgn2_6/ptodata/1/pubna/US08_NEW_PUB.seq.*
- 8: /cgn2_6/ptodata/1/pubna/US08_PUBCOMB.seq.*
- 9: /cgn2_6/ptodata/1/pubna/US09_NEW_PUB.seq.*
- 10: /cgn2_6/ptodata/1/pubna/US09_PUBCOMB.seq.*
- 11: /cgn2_6/ptodata/1/pubna/US10_NEW_PUB.seq.*
- 12: /cgn2_6/ptodata/1/pubna/US10_PUBCOMB.seq.*
- 13: /cgn2_6/ptodata/1/pubna/US60_NEW_PUB.seq.*
- 14: /cgn2_6/ptodata/1/pubna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	21	100.0	1740	10	US-09-735-705-347
2	21	100.0	1740	10	US-09-850-716A-347
3	21	100.0	1740	10	US-09-897-778-347
4	21	100.0	1743	10	US-09-897-778-447
5	21	100.0	1743	10	US-09-897-778-450
6	21	100.0	1764	10	US-09-850-716A-428
7	21	100.0	1764	10	US-09-897-778-428
8	21	100.0	4159	10	US-09-899-651-4
9	21	100.0	4181	10	US-09-735-705-175
10	21	100.0	4181	10	US-09-954-456-715
11	21	100.0	4181	10	US-09-850-716A-175
12	21	100.0	4181	10	US-09-897-778-175
13	16.2	77.1	359	10	US-09-864-761-23958
14	16.2	77.1	476	10	US-09-864-761-1856
15	16.2	77.1	479	10	US-09-880-107-2877
16	16.2	77.1	597	10	US-09-864-761-7227
17	16.2	77.1	1178	10	US-09-925-300-662
18	16.2	77.1	9962	9	US-09-764-872-797
19	15.8	75.2	32191	10	US-09-764-864-1678

c	20	15.8	75.2	1691139	9	US-10-067-514-1	Sequence 1, Appli
	21	15.4	73.3	533	10	US-09-867-701-4800	Sequence 4800, Ap
	22	15.4	73.3	626	10	US-09-770-149-734	Sequence 734, App
	23	15.4	73.3	3555	12	US-10-044-090-527	Sequence 527, App
	24	15.4	73.3	5866	12	US-10-044-090-215	Sequence 215, App
	25	15.2	72.4	1686	10	US-09-731-872-183	Sequence 183, App
	26	15.2	72.4	3468	9	US-10-042-141-36	Sequence 36, Appl
	27	15.2	72.4	22161	9	US-09-726-643-36	Sequence 36, Appl
	28	15.2	72.4	22161	9	US-10-092-154-1020	Sequence 1020, Ap
	29	15.2	72.4	106323	10	US-09-764-847-1020	Sequence 1020, Ap
	30	15.2	72.4	536165	9	US-09-803-661-3	Sequence 3, Appli
	31	15.2	72.4	536165	9	US-09-939-964-1	Sequence 113, App
	32	15	71.4	373	9	US-09-933-797-113	Sequence 113, App
	33	14.8	70.5	167	10	US-09-728-445-394	Sequence 394, App
	34	14.8	70.5	275	10	US-09-783-590-3645	Sequence 3645, Ap
	35	14.8	70.5	465	10	US-09-864-761-1651	Sequence 1651, Ap
	36	14.8	70.5	487	10	US-09-864-761-2009	Sequence 2009, Ap
	37	14.8	70.5	501	10	US-09-998-598-802	Sequence 802, App
	38	14.8	70.5	549	10	US-09-864-761-9819	Sequence 9819, Ap
	39	14.8	70.5	601	10	US-09-920-300A-908	Sequence 908, App
	40	14.8	70.5	601	12	US-10-033-528-908	Sequence 908, App
	41	14.8	70.5	1074	9	US-09-938-842A-2656	Sequence 2656, Ap
	42	14.8	70.5	3180	10	US-09-949-192-14	Sequence 14, Appl
	43	14.8	70.5	4322	9	US-10-098-841-292	Sequence 292, Appl
	44	14.8	70.5	18355	7	US-08-781-986A-67	Sequence 67, Appl
	45	14.8	70.5	20029	9	US-10-092-154-1624	Sequence 1624, Ap

ALIGNMENTS

RESULT 1
US-09-735-705-347
; Sequence 347, Application US/09735705
; Patent No. US20020052329A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Fan, Liqun
; APPLICANT: Kalos, Michael D.
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Hosken, Nancy
; APPLICANT: Fanger, Gary R.
; APPLICANT: Li, Samuel X.
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Henderson, Robert A.
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Fanger, Neil
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.455C14
; CURRENT APPLICATION NUMBER: US/09/735,705
; CURRENT FILING DATE: 2000-12-12
; NUMBER OF SEQ ID NOS: 419
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 347
; LENGTH: 1740
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-735-705-347

Query Match 100.0%; Score 21; DB 10; Length 1740;
Best Local Similarity 100.0%; Pred. No. 0.28; 0; Indels 0; Gaps 0;
Matches 21; Conservative 0; Mismatches 0;

Qy 1 GAAAGTATCTTCAAGGAGCC 21

Db 55 GAAAGTATCTTCAAGGAGCC 75

RESULT 2

US-09-850-716A-347
; Sequence 347, Application US/09850716A

us-09-270-437d-9.rnpb

Mon Apr 21 10:30:54 2003

```

; Patent No. US20020115139A1
; GENERAL INFORMATION:
; APPLICANT: Kalos, Michael D.
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Retter, Marc W.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.455C15
; CURRENT FILING DATE: 2001-05-07
; NUMBER OF SEQ ID NOS: 440
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 347
; LENGTH: 1740
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-850-716A-347

Query Match      100.0%; Score 21; DB 10; Length 1740;
Best Local Similarity 100.0%; Pred. No. 0.28;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GAAAGTATCTTCAAGGAGCC 21
Db 55 GAAAGTATCTTCAAGGAGCC 75

RESULT 3
US-09-897-778-347
; Sequence 347, Application US/09897778
; Patent No. US20020147143A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Marnerakis, Margarita
; APPLICANT: Fanger, Gary R.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Watanabe, Yoshihiro
; APPLICANT: Henderson, Robert A.
; APPLICANT: Peckham, David W.
; APPLICANT: Fanger, Neil
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.455C16
; CURRENT FILING DATE: 2001-06-28
; NUMBER OF SEQ ID NOS: 467
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 347
; LENGTH: 1740
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-897-778-347

Query Match      100.0%; Score 21; DB 10; Length 1740;
Best Local Similarity 100.0%; Pred. No. 0.28;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GAAAGTATCTTCAAGGAGCC 21
Db 55 GAAAGTATCTTCAAGGAGCC 75

RESULT 4
US-09-897-778-447
; Sequence 447, Application US/09897778
; Patent No. US20020147143A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Marnerakis, Margarita
; APPLICANT: Fanger, Gary R.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Watanabe, Yoshihiro
; APPLICANT: Henderson, Robert A.
; APPLICANT: Peckham, David W.
; APPLICANT: Fanger, Neil
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.455C16
; CURRENT FILING DATE: 2001-06-28
; NUMBER OF SEQ ID NOS: 467
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 447
; LENGTH: 1743
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-897-778-447

Query Match      100.0%; Score 21; DB 10; Length 1743;
Best Local Similarity 100.0%; Pred. No. 0.28;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GAAAGTATCTTCAAGGAGCC 21
Db 55 GAAAGTATCTTCAAGGAGCC 75

RESULT 5
US-09-897-778-450
; Sequence 450, Application US/09897778
; Patent No. US20020147143A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Marnerakis, Margarita
; APPLICANT: Fanger, Gary R.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Watanabe, Yoshihiro
; APPLICANT: Henderson, Robert A.
; APPLICANT: Peckham, David W.
; APPLICANT: Fanger, Neil
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.455C16
; CURRENT FILING DATE: 2001-06-28
; NUMBER OF SEQ ID NOS: 467
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 450
; LENGTH: 1743
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-897-778-450

Query Match      100.0%; Score 21; DB 10; Length 1743;
Best Local Similarity 100.0%; Pred. No. 0.28;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GAAAGTATCTTCAAGGAGCC 21
Db 55 GAAAGTATCTTCAAGGAGCC 75

RESULT 6
US-09-850-716A-428
; Sequence 428, Application US/09850716A
; Patent No. US20020115139A1
; GENERAL INFORMATION:
; APPLICANT: Kalos, Michael D.
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Retter, Marc W.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.455C15

```


; CURRENT APPLICATION NUMBER: US/09/850,716A
; PRIOR FILING DATE: 2001-05-07
; NUMBER OF SEQ ID NOS: 440
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 428
; LENGTH: 1764
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-850-716A-428

Query Match 100.0%; Score 21; DB 10; Length 1764;
Best Local Similarity 100.0%; Pred. No. 0.28; Indels 0; Gaps 0;
Matches 21; Conservative 0; Mismatches 0;

QY 1 GAAAGTATCTTCAAGGACGCC 21
| | | | | | | | | | | | | | | | | | | | | |
DB 76 GAAAGTATCTTCAAGGACGCC 96

RESULT 7

US-09-897-778-428
; Sequence 428, Application US/09897778
; Patent No. US20020147143A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Marnerakis, Margarita
; APPLICANT: Fanger, Gary R.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Watanabe, Yoshihiro
; APPLICANT: Henderson, Robert A.
; APPLICANT: Peckham, David W.
; APPLICANT: Fanger, Neil
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.455C16
; CURRENT APPLICATION NUMBER: US/09/897,778
; CURRENT FILING DATE: 2001-06-28
; NUMBER OF SEQ ID NOS: 467
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 428
; LENGTH: 1764
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-897-778-428

Query Match 100.0%; Score 21; DB 10; Length 1764;
Best Local Similarity 100.0%; Pred. No. 0.28; Indels 0; Gaps 0;
Matches 21; Conservative 0; Mismatches 0;

QY 1 GAAAGTATCTTCAAGGACGCC 21
| | | | | | | | | | | | | | | | | | | | | |
DB 76 GAAAGTATCTTCAAGGACGCC 96

RESULT 8

US-09-899-651-4
; Sequence 4, Application US/09899651
; Patent No. US20020111470A1
; GENERAL INFORMATION:
; APPLICANT: Chen, Yao-Tseng
; APPLICANT: Gure, Ali
; APPLICANT: Tsang, Solam
; APPLICANT: Stockert, Elisabeth
; APPLICANT: Jager, Elke
; APPLICANT: Knuth, Alexander
; APPLICANT: Old, Lloyd J.
; TITLE OF INVENTION: Isolated Nucleic Acid Molecules Encoding Cancer
; FILE REFERENCE: Associated
; TITLE OF INVENTION: Antigen, The Antigens Per Se, And Uses Thereof
; FILE REFERENCE: LUD 5538
; CURRENT APPLICATION NUMBER: US/09/899,651
; CURRENT FILING DATE: 2001-07-06

; PRIOR APPLICATION NUMBER: US/09/061,709
; PRIOR FILING DATE: 1998-04-17
; NUMBER OF SEQ ID NOS: 8
; SEQ ID NO 4
; LENGTH: 4159
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
US-09-899-651-4

Query Match 100.0%; Score 21; DB 10; Length 4159;
Best Local Similarity 100.0%; Pred. No. 0.32; Indels 0; Gaps 0;
Matches 21; Conservative 0; Mismatches 0;

QY 1 GAAAGTATCTTCAAGGACGCC 21
| | | | | | | | | | | | | | | | | | | | | |
DB 305 GAAAGTATCTTCAAGGACGCC 325

RESULT 9

US-09-735-705-175
; Sequence 175, Application US/09735705
; Patent No. US20020052329A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Fan, Liqun
; APPLICANT: Kalos, Michael D.
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Hosken, Nancy
; APPLICANT: Fanger, Gary R.
; APPLICANT: Li, Samuel X.
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Henderson, Robert A.
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Fanger, Neil
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.455C14
; CURRENT APPLICATION NUMBER: US/09/735,705
; CURRENT FILING DATE: 2000-12-12
; NUMBER OF SEQ ID NOS: 419
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 175
; LENGTH: 4181
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (3347)
; OTHER INFORMATION: n=A,T,C or G
; NAME/KEY: unsure
; LOCATION: (3502)
; OTHER INFORMATION: n=A,T,C or G
; NAME/KEY: unsure
; LOCATION: (3506)
; OTHER INFORMATION: n=A,T,C or G
; NAME/KEY: unsure
; LOCATION: (3520)
; OTHER INFORMATION: n=A,T,C or G
; NAME/KEY: unsure
; LOCATION: (3538)
; OTHER INFORMATION: n=A,T,C or G
; NAME/KEY: unsure
; LOCATION: (3549)
; OTHER INFORMATION: n=A,T,C or G
; NAME/KEY: unsure
; LOCATION: (3646)
; OTHER INFORMATION: n=A,T,C or G
; NAME/KEY: unsure
; LOCATION: (3940)
; OTHER INFORMATION: n=A,T,C or G
; NAME/KEY: unsure

[illegible]

; OTHER INFORMATION: n=A,T,C or G
 ; NAME/KEY: unsure
 ; LOCATION: (4088)
 ; OTHER INFORMATION: n=A,T,C or G
 ; NAME/KEY: unsure
 ; LOCATION: (4115)
 ; OTHER INFORMATION: n=A,T,C or G
 US-09-850-716A-175

Query Match 100.0%; Score 21; DB 10; Length 4181;
 Best Local Similarity 100.0%; Pred. No. 0.32; Mismatches 0; Indels 0; Gaps 0;
 Matches 21; Conservative 0;

Qy 1 GAAAGTATCTTCAAGGACGCC 21
 |||||
 Db 305 GAAAGTATCTTCAAGGACGCC 325

RESULT 12

US-09-897-778-175
 ; Sequence 175, Application US/09897778
 ; Patent No. US20020147143A1
 ; GENERAL INFORMATION:

; APPLICANT: Wang, Tonglong
 ; APPLICANT: Marnerakis, Margarita
 ; APPLICANT: Fanger, Gary R.
 ; APPLICANT: Vedvick, Thomas S.
 ; APPLICANT: Carter, Darrick
 ; APPLICANT: Watanabe, Yoshihiro
 ; APPLICANT: Henderson, Robert A.
 ; APPLICANT: Peckham, David W.
 ; APPLICANT: Fanger, Neil

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
 ; FILE REFERENCE: 210121.455C16
 ; CURRENT APPLICATION NUMBER: US/09/897,778

; NUMBER OF SEQ ID NOS: 467
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 175
 ; LENGTH: 4181
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens

; FEATURE:
 ; NAME/KEY: misc_feature
 ; LOCATION: 3347, 3502, 3506, 3520, 3538, 3549, 3646, 3940, 3968, 3974,
 ; LOCATION: 4036, 4056, 4062, 4080, 4088, 4115
 ; OTHER INFORMATION: n = A,T,C or G
 US-09-897-778-175

Query Match 100.0%; Score 21; DB 10; Length 4181;
 Best Local Similarity 100.0%; Pred. No. 0.32; Mismatches 0; Indels 0; Gaps 0;
 Matches 21; Conservative 0;

Qy 1 GAAAGTATCTTCAAGGACGCC 21
 |||||
 Db 305 GAAAGTATCTTCAAGGACGCC 325

RESULT 13

US-09-864-761-23958/c
 ; Sequence 23958, Application US/09864761
 ; Patent No. US20020048763A1
 ; GENERAL INFORMATION:

; APPLICANT: Penn, Sharron G.
 ; APPLICANT: Rank, David R.
 ; APPLICANT: Hanzel, David K.
 ; APPLICANT: Chen, Wensheng

; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
 ; FILE REFERENCE: Aeonica-X-1
 ; CURRENT APPLICATION NUMBER: US/09/864,761
 ; CURRENT FILING DATE: 2001-05-23

; PRIOR APPLICATION NUMBER: US 60/180,312
 ; PRIOR FILING DATE: 2000-02-04
 ; PRIOR APPLICATION NUMBER: US 60/207,456
 ; PRIOR FILING DATE: 2000-05-26
 ; PRIOR APPLICATION NUMBER: US 09/632,366
 ; PRIOR FILING DATE: 2000-08-03
 ; PRIOR APPLICATION NUMBER: GB 24263.6
 ; PRIOR FILING DATE: 2000-10-04
 ; PRIOR APPLICATION NUMBER: US 60/236,359
 ; PRIOR FILING DATE: 2000-09-27
 ; PRIOR APPLICATION NUMBER: PCT/US01/00666
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00667
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00664
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00669
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00665
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00668
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00663
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00662
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00661
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00670
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: US 60/234,687
 ; PRIOR FILING DATE: 2000-09-21
 ; PRIOR APPLICATION NUMBER: US 09/608,408
 ; PRIOR FILING DATE: 2000-06-30
 ; PRIOR APPLICATION NUMBER: US 09/774,203
 ; PRIOR FILING DATE: 2001-01-29
 ; NUMBER OF SEQ ID NOS: 49117
 ; SOFTWARE: Annonax Sequence Listing Engine vers. 1.1
 ; SEQ ID NO 23958
 ; LENGTH: 359
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; OTHER INFORMATION: MAP TO AL121777.16
 ; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 2.9
 ; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.2
 ; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 0.96
 ; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.4
 ; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 2
 ; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.5
 ; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.4
 ; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.8
 ; OTHER INFORMATION: NT HIT: gi4507208, EVALUE 1.00e-15
 ; OTHER INFORMATION: SWISSPROT HIT: P19623, EVALUE 2.00e-14
 ; OTHER INFORMATION: EST_HUMAN HIT: BE904452.1, EVALUE 2.00e-15
 US-09-864-761-23958

Query Match 77.1%; Score 16.2; DB 10; Length 359;
 Best Local Similarity 85.7%; Pred. No. 65;
 Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 GAAAGTATCTTCAAGGACGCC 21
 |||||
 Db 279 GAAAGCTCTTCAAGGACGCC 259

RESULT 14

US-09-864-761-1856
 ; Sequence 1856, Application US/09864761
 ; Patent No. US20020048763A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Penn, Sharron G.
 ; APPLICANT: Rank, David R.

US-09-880-107-2877
; Sequence 2877, Application US/09880107
; Patent No. US20020142981A1
; GENERAL INFORMATION:
; APPLICANT: Horne, Darci T.
; APPLICANT: Vockley, Joseph G.
; APPLICANT: Scherf, Uwe
; APPLICANT: Gene Logic, Inc.
; TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
; FILE REFERENCE: 44921-5028-WO
; CURRENT APPLICATION NUMBER: US/09/880,107
; CURRENT FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: US 60/211,379
; PRIOR FILING DATE: 2000-06-14
; PRIOR APPLICATION NUMBER: US 60/237,054
; PRIOR FILING DATE: 2000-10-02
; NUMBER OF SEQ ID NOS: 3950
; SOFTWARE: Patentin ver. 2.1
; SEQ ID NO 2877
; LENGTH: 479
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20020142981A1 R49047
US-09-880-107-2877

Query Match 77.1%; Score 16.2; DB 10; Length 479;
Best Local Similarity 85.7%; Pred. No. 69;
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 GAAAGTATCTTCAAGGAGGCC 21
|||||
DB 223 GAAAGGAACTTCAAGGAGGCC 243
|||||

Search completed: April 18, 2003, 10:13:04
Job time : 92.3485 secs

US-09-864-761-1856
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
; FILE REFERENCE: Aeonica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 1856
; LENGTH: 476
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO Z95330.7
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.4
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.4
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.4
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.3
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.6
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 4
; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 4.1
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.3
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 2.1
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.3
US-09-864-761-1856

Query Match 77.1%; Score 16.2; DB 10; Length 476;
Best Local Similarity 85.7%; Pred. No. 68;
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 GAAAGTATCTTCAAGGAGGCC 21
|||||
DB 428 GAAAGGCTTCTCAAGGAGGCC 448
|||||

RESULT 15

Result No.	Score	Query		DB	ID	Description
		Match	Length			
1	21	100.0	422	16	US-09-271-490-13091	Sequence 13091, A
2	21	100.0	422	34	US-09-925-555-13091	Sequence 13091, A
3	21	100.0	422	38	US-10-032-354-13091	Sequence 13091, A
4	21	100.0	428	16	US-09-287-618-27259	Sequence 27259, A
5	21	100.0	529	17	US-09-399-932-638	Sequence 638, App
6	21	100.0	556	25	US-09-652-123-4489	Sequence 4489, App
7	21	100.0	834	25	US-09-652-123-8442	Sequence 8442, Ap
8	21	100.0	843	17	US-09-399-932-4493	Sequence 4493, Ap
9	21	100.0	1740	1	PCT-US01-47576-347	Sequence 347, App
10	21	100.0	1740	21	US-09-542-615A-347	Sequence 347, App
11	21	100.0	1740	23	US-09-606-421A-347	Sequence 347, App
12	21	100.0	1740	23	US-09-606-421B-347	Sequence 347, App
13	21	100.0	1740	24	US-09-630-940A-347	Sequence 347, App
14	21	100.0	1740	24	US-09-630-940B-347	Sequence 347, App
15	21	100.0	1740	26	US-09-662-786-347	Sequence 347, App
16	21	100.0	1740	27	US-09-685-696-347	Sequence 347, App
17	21	100.0	1740	29	US-09-735-705-347	Sequence 347, App
18	21	100.0	1740	32	US-09-850-716-347	Sequence 347, App
19	21	100.0	1740	32	US-09-850-716A-347	Sequence 347, App
20	21	100.0	1740	33	US-09-897-778-347	Sequence 347, App
21	21	100.0	1740	38	US-10-007-700-347	Sequence 347, App

22	21	100.0	1740	40	US-10-117-982-347	Sequence 347, App
23	21	100.0	1740	40	US-10-117-982-478	Sequence 478, App
24	21	100.0	1740	40	US-10-117-982-483	Sequence 483, App
25	21	100.0	1740	1	PCT-US01-47576-447	Sequence 447, App
26	21	100.0	1743	1	PCT-US01-47576-447	Sequence 450, App
27	21	100.0	1743	33	US-09-897-778-447	Sequence 447, App
28	21	100.0	1743	33	US-09-897-778-450	Sequence 450, App
29	21	100.0	1743	38	US-10-007-700-447	Sequence 447, App
30	21	100.0	1743	38	US-10-007-700-450	Sequence 450, App
31	21	100.0	1743	40	US-10-117-982-447	Sequence 447, App
32	21	100.0	1743	40	US-10-117-982-450	Sequence 447, App
33	21	100.0	1764	1	PCT-US01-47576-428	Sequence 428, App
34	21	100.0	1764	32	US-09-850-716-428	Sequence 428, App
35	21	100.0	1764	32	US-09-850-716-428	Sequence 428, App
36	21	100.0	1764	33	US-09-897-778-428	Sequence 428, App
37	21	100.0	1764	38	US-10-007-700-428	Sequence 428, App
38	21	100.0	1764	40	US-10-117-982-428	Sequence 428, App
39	21	100.0	4159	16	US-09-270-437-4	Sequence 4, Appli
40	21	100.0	4159	33	US-09-899-651-4	Sequence 4, Appli
41	21	100.0	4181	1	PCT-US01-47576-175	Sequence 175, App
42	21	100.0	4181	18	US-09-466-396A-175	Sequence 175, App
43	21	100.0	4181	18	US-09-476-496A-175	Sequence 175, App
44	21	100.0	4181	18	US-09-480-884A-175	Sequence 175, App
45	21	100.0	4191	19	US-09-510-376A-175	Sequence 175, App

ATTACHMENTS

```

RESULT 1
US-09-271-490-13091
; Sequence 13091, Application US/09271490
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; FROM VARIOUS CDNA LIBRARIES
; FILE REFERENCE: 20411-767
; CURRENT APPLICATION NUMBER: US/09/271,490
; CURRENT FILING DATE: 1999-03-18
; NUMBER OF SEQ ID NOS: 19424
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 13091
; LENGTH: 422
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(422)
; OTHER INFORMATION: n = A,T,C or G
; US-09-271-490-13091

```

Query Match	100.0%;	Score 21;	DB 16;	Length 422;
Best Local Similarity	100.0%;	Pred. No. 3.9;		
		Mismatches	0.	Indels
			0.	Gaps

```

QY      1 GAAAGTATCTTCACGAGCC 21
        |||||
Db     369 GAAAGTATCTTCACGAGCC 389

```

RESULT 2
US-09-925-552-13091
; Sequence 13091, Application US/09925552
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; FROM VARIOUS CDNA LIBRARIES
; FILE REFERENCE: 20411-767
; CURRENT APPLICATION NUMBER: US/09/925,552
; CURRENT FILING DATE: 2001-08-09
; PRIOR APPLICATION NUMBER: 09/271,490
; PRIOR FILING DATE: 1999-03-18
; NUMBER OF SEQ ID NOS: 19424

```

; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 13091
; LENGTH: 422
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(422)
; OTHER INFORMATION: n = A,T,C or G
US-09-925-552-13091

```

Query Match	100.0%;	Score 21;	DB 34;	Length 422;
Best Local Similarity	100.0%;	Pred. No. 3.9;		
Matches 21; Conservative	0;	Mismatches 0;	Indels 0;	Gaps 0;

Qy 1 GAAAGTATCTTCAAGGACGCC 21
|||
nb 369 GAAAGTATCTTCAAGGACGCC 389

```

RESULT 3
US-10-032-354-13091
Sequence 13091, Application US/10032354
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; FROM VARIOUS CDNA LIBRARIES
; FILE REFERENCE: 20411-767
; CURRENT APPLICATION NUMBER: US/10/032,354
; CURRENT FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: 09/271,490
; PRIOR FILING DATE: 1999-03-18
; NUMBER OF SEQ ID NOS: 19424
; SOFTWARE: Fast-SEQ for Windows Version 3.0
; SEQ ID NO 13091
; LENGTH: 422
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(422)
; OTHER INFORMATION: n = A,T,C or G
; vs_10_032-354-13091

```

Query Match	100.0%;	Score 21;	DB 38;	Length 422;
Best Local Similarity	100.0%;	Pred. No. 3.9;		
Matches 21;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

Qy 1 GAAAGTATCTTCAAGGACGCC 21
|||
nb 369 GAAAGTATCTTCAAGGACGCC 389

```

RESULT 4
US-09-287-618-27259
; Sequence 27259, Application US/09287618
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; FROM VARIOUS CDNA LIBRARIES
; FILE REFERENCE: 20411-768
; CURRENT APPLICATION NUMBER: US/09/287,618
; CURRENT FILING DATE: 1999-04-02
; NUMBER OF SEQ ID NOS: 35865
; SOFTWARE: Fast-Seq for Windows Version 3.0
; SEQ ID NO 27259
; LENGTH: 438
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-287-618-27259

```

Query Match	100.0%;	Score 21;	DB 16;	Length 438;
Best Local Similarity	100.0%;	Pred. No. 3.9;		

Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAAAGTATCTTCAAGGACGCC 21
|||||
Db 388 GAAAGTATCTTCAAGGACGCC 408

RESULT 5

US-09-399-932-638
; Sequence 638, Application US/09399932
; GENERAL INFORMATION:
; APPLICANT: Holtzman, Douglas A.
; APPLICANT: Gearing, David P.
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES DERIVED FROM A
; TITLE OF INVENTION: HUMAN BURKITT'S LYMPHOMA LIBRARY
; FILE REFERENCE: MLN98-47PM
; CURRENT APPLICATION NUMBER: US/09/399,932
; CURRENT FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: 60/100,464
; PRIOR FILING DATE: 1998-09-15
; PRIOR APPLICATION NUMBER: 60/101,670
; PRIOR FILING DATE: 1998-09-24
; PRIOR APPLICATION NUMBER: 60/106,456
; PRIOR FILING DATE: 1998-10-30
; PRIOR APPLICATION NUMBER: 60/126,902
; PRIOR FILING DATE: 1999-03-30
; PRIOR APPLICATION NUMBER: 60/132,067
; PRIOR FILING DATE: 1999-04-30
; PRIOR APPLICATION NUMBER: 60/144,447
; PRIOR FILING DATE: 1999-07-16
; NUMBER OF SEQ ID NOS: 5743
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 638
; LENGTH: 529
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(529)
; OTHER INFORMATION: n = A,T,C or G
US-09-399-932-638

Query Match 100.0%; Score 21; DB 17; Length 529;
Best Local Similarity 100.0%; Pred. No. 4.1;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAAAGTATCTTCAAGGACGCC 21
|||||
Db 94 GAAAGTATCTTCAAGGACGCC 114

RESULT 6

US-09-652-123-4489
; Sequence 4489, Application US/09652123
; GENERAL INFORMATION:
; APPLICANT: Holtzman, Douglas A.
; APPLICANT: Gearing, David P.
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES DERIVED FROM A
; TITLE OF INVENTION: HUMAN BURKITT'S LYMPHOMA LIBRARY
; FILE REFERENCE: 1600.1186-001
; CURRENT APPLICATION NUMBER: US/09/652,123
; CURRENT FILING DATE: 2000-08-30
; PRIOR APPLICATION NUMBER: 60/151,135
; PRIOR FILING DATE: 1999-08-30
; NUMBER OF SEQ ID NOS: 9796
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4489
; LENGTH: 556
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(556)
; OTHER INFORMATION: n = A,T,C or G

US-09-652-123-4489

Query Match 100.0%; Score 21; DB 25; Length 556;
Best Local Similarity 100.0%; Pred. No. 4.1;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAAAGTATCTTCAAGGACGCC 21
|||||
Db 409 GAAAGTATCTTCAAGGACGCC 429

RESULT 7

US-09-652-123-8442
; Sequence 8442, Application US/09652123
; GENERAL INFORMATION:
; APPLICANT: Holtzman, Douglas A.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID MOLECULES AND USES
; TITLE OF INVENTION: THEREFOR
; FILE REFERENCE: 1600.1186-001
; CURRENT APPLICATION NUMBER: US/09/652,123
; CURRENT FILING DATE: 2000-08-30
; PRIOR APPLICATION NUMBER: 60/151,135
; PRIOR FILING DATE: 1999-08-30
; NUMBER OF SEQ ID NOS: 9796
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8442
; LENGTH: 834
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(834)
; OTHER INFORMATION: n = A,T,C or G
US-09-652-123-8442

Query Match 100.0%; Score 21; DB 25; Length 834;
Best Local Similarity 100.0%; Pred. No. 4.5;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAAAGTATCTTCAAGGACGCC 21
|||||
Db 407 GAAAGTATCTTCAAGGACGCC 427

RESULT 8

US-09-399-932-4493
; Sequence 4493, Application US/09399932
; GENERAL INFORMATION:
; APPLICANT: Holtzman, Douglas A.
; APPLICANT: Gearing, David P.
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES DERIVED FROM A
; TITLE OF INVENTION: HUMAN BURKITT'S LYMPHOMA LIBRARY
; FILE REFERENCE: MLN98-47PM
; CURRENT APPLICATION NUMBER: US/09/399,932
; CURRENT FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: 60/100,464
; PRIOR FILING DATE: 1998-09-15
; PRIOR APPLICATION NUMBER: 60/101,670
; PRIOR FILING DATE: 1998-09-24
; PRIOR APPLICATION NUMBER: 60/106,456
; PRIOR FILING DATE: 1998-10-30
; PRIOR APPLICATION NUMBER: 60/126,902
; PRIOR FILING DATE: 1999-03-30
; PRIOR APPLICATION NUMBER: 60/132,067
; PRIOR FILING DATE: 1999-04-30
; PRIOR APPLICATION NUMBER: 60/144,447
; PRIOR FILING DATE: 1999-07-16
; NUMBER OF SEQ ID NOS: 5743
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4493
; LENGTH: 843
; TYPE: DNA
; ORGANISM: Homo sapiens

US-09-606-421B-347

Query Match 100.0%; Score 21; DB 23; Length 1740;
Best Local Similarity 100.0%; Pred. No. 5.2; Indels 0; Gaps 0;
Matches 21; Conservative 0; Mismatches 0;

QY 1 GAAAGTATCTTCAAGGACGCC 21
|||||
Db 55 GAAAGTATCTTCAAGGACGCC 75

RESULT 13

US-09-630-940A-347
; Sequence 347, Application US/09630940A

; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Fan, Liqun
; APPLICANT: Kalos, Michael D.
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Hosken, Nancy
; APPLICANT: Fanger, Gary R.
; APPLICANT: Li, Samuel X.
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Henderson, Robert A.
; APPLICANT: McNeill, Patricia D.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.455C10
; CURRENT APPLICATION NUMBER: US/09/630,940A
; CURRENT FILING DATE: 2000-08-02
; NUMBER OF SEQ ID NOS: 367
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 347
; LENGTH: 1740
; TYPE: DNA

US-09-630-940A-347
; ORGANISM: Homo sapiens

Query Match 100.0%; Score 21; DB 24; Length 1740;
Best Local Similarity 100.0%; Pred. No. 5.2; Indels 0; Gaps 0;
Matches 21; Conservative 0; Mismatches 0;

QY 1 GAAAGTATCTTCAAGGACGCC 21
|||||
Db 55 GAAAGTATCTTCAAGGACGCC 75

RESULT 14

US-09-630-940B-347
; Sequence 347, Application US/09630940B

; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Fan, Liqun
; APPLICANT: Kalos, Michael D.
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Hosken, Nancy
; APPLICANT: Fanger, Gary R.
; APPLICANT: Li, Samuel X.
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Henderson, Robert A.
; APPLICANT: McNeill, Patricia D.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.455C10
; CURRENT APPLICATION NUMBER: US/09/630,940B
; CURRENT FILING DATE: 2000-08-02
; NUMBER OF SEQ ID NOS: 367
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 347
; LENGTH: 1740
; TYPE: DNA

; ORGANISM: Homo sapiens
US-09-630-940B-347

Query Match 100.0%; Score 21; DB 24; Length 1740;
Best Local Similarity 100.0%; Pred. No. 5.2; Indels 0; Gaps 0;
Matches 21; Conservative 0; Mismatches 0;

QY 1 GAAAGTATCTTCAAGGACGCC 21
|||||
Db 55 GAAAGTATCTTCAAGGACGCC 75

RESULT 15

US-09-662-786-347
; Sequence 347, Application US/09662786

; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Fan, Liqun
; APPLICANT: Kalos, Michael D.
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Hosken, Nancy
; APPLICANT: Fanger, Gary R.
; APPLICANT: Li, Samuel X.
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Henderson, Robert A.
; APPLICANT: McNeill, Patricia D.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.455C12
; CURRENT APPLICATION NUMBER: US/09/662,786
; CURRENT FILING DATE: 2000-09-15
; NUMBER OF SEQ ID NOS: 381
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 347
; LENGTH: 1740
; TYPE: DNA

US-09-662-786-347
; ORGANISM: Homo sapiens

Query Match 100.0%; Score 21; DB 26; Length 1740;
Best Local Similarity 100.0%; Pred. No. 5.2; Indels 0; Gaps 0;
Matches 21; Conservative 0; Mismatches 0;

QY 1 GAAAGTATCTTCAAGGACGCC 21
|||||
Db 55 GAAAGTATCTTCAAGGACGCC 75

Search completed: April 18, 2003, 09:35:40
Job time : 1159.95 secs

GenCore version 5.1.4 p5 4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 18, 2003, 06:35:28 ; Search time 234.023 Seconds
(without alignments)
400.770 Million cell updates/sec

Title: US-09-270-437D-9

Perfect score: 21
Sequence: 1 gaaagtattctcaaggagcc 21

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 5897297 seqs, 223308081 residues

Total number of hits satisfying chosen parameters: 11794594

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Pending Patents NA New:

- 1: /cgn2_6/ptodata/1/pna/PCT_NEW_COMB.seq.*
- 2: /cgn2_6/ptodata/1/pna/US06_NEW_COMB.seq.*
- 3: /cgn2_6/ptodata/1/pna/US07_NEW_COMB.seq.*
- 4: /cgn2_6/ptodata/1/pna/US08_NEW_COMB.seq.*
- 5: /cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq.*
- 6: /cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq2.*
- 7: /cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq3.*
- 8: /cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq.*
- 9: /cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq2.*
- 10: /cgn2_6/ptodata/1/pna/US60_NEW_COMB.seq.*
- 11: /cgn2_6/ptodata/1/pna/US60_NEW_COMB.seq2.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	21	100.0	21	5	US-09-270-437D-9
2	21	100.0	861	6	US-09-724-676-44306
3	21	100.0	861	6	US-09-724-676A-44306
4	21	100.0	1558	8	US-10-170-235-4888
5	21	100.0	1740	9	US-10-313-986-347
6	21	100.0	1740	9	US-10-313-986-478
7	21	100.0	1740	9	US-10-313-986-483
8	21	100.0	1743	9	US-10-313-986-447
9	21	100.0	1743	9	US-10-313-986-450
10	21	100.0	1764	9	US-10-313-986-428
11	21	100.0	1789	9	US-10-313-986-485
12	21	100.0	2137	6	US-09-724-676-44307
13	21	100.0	2137	6	US-09-724-676A-44307
14	21	100.0	4060	6	US-09-724-676-44305
15	21	100.0	4060	6	US-09-724-676A-44305
16	21	100.0	4159	5	US-09-270-437D-4
17	21	100.0	4181	9	US-10-313-986-175
18	21	100.0	4181	9	US-10-313-986-119-55
19	21	100.0	34555	9	US-10-313-986-479
20	19	90.5	201	11	US-60-453-135-15387
21	19	90.5	201	11	US-60-453-050-15387
22	19	90.5	3730	11	US-60-453-135-1

23	19	90.5	3730	11	US-60-453-050-1	Sequence 1, Appli
24	18.4	87.6	1748	9	US-10-144-771-21075	Sequence 21075, A
25	17.4	82.9	25	11	US-60-427-808-455974	Sequence 455974, A
26	17.4	82.9	25	11	US-60-427-808-455975	Sequence 455975, A
27	16.8	80.0	268	6	US-09-532-315B-13848	Sequence 13848, A
28	16.8	80.0	414	7	US-09-675-784A-4706	Sequence 4706, Ap
29	16.8	80.0	7681	9	US-10-144-771-11275	Sequence 11275, A
30	16.2	77.1	79	6	US-09-532-315B-9552	Sequence 9552, Ap
31	16.2	77.1	123	6	US-09-532-315B-9557	Sequence 9557, Ap
32	16.2	77.1	201	11	US-60-452-680-84031	Sequence 84031, A
33	16.2	77.1	201	11	US-60-453-135-55791	Sequence 55791, A
34	16.2	77.1	201	11	US-60-453-050-55791	Sequence 55791, A
35	16.2	77.1	220	9	US-10-349-781-19919	Sequence 19919, A
36	16.2	77.1	254	6	US-09-532-315B-13854	Sequence 13854, A
37	16.2	77.1	272	6	US-09-532-315B-9553	Sequence 9553, Ap
38	16.2	77.1	300	9	US-10-144-771-39459	Sequence 39459, A
39	16.2	77.1	325	6	US-09-532-315B-9544	Sequence 9544, Ap
40	16.2	77.1	476	9	US-10-203-138A-1905	Sequence 1905, Ap
41	16.2	77.1	479	8	US-10-401-223-7878	Sequence 7878, Ap
42	16.2	77.1	538	6	US-09-532-315B-9551	Sequence 9551, Ap
43	16.2	77.1	544	6	US-09-532-315B-9556	Sequence 9556, Ap
44	16.2	77.1	568	6	US-09-532-315B-9555	Sequence 9555, Ap
45	16.2	77.1	601	7	US-09-947-907-19977	Sequence 19977, A

ALIGNMENTS

RESULT 1
US-09-270-437D-9
; Sequence 9, Application US/09270437D
; GENERAL INFORMATION:
; APPLICANT: Chen, Yao-Tseung
; APPLICANT: Gure, Ali
; APPLICANT: Tsang, Solam
; APPLICANT: Stockert, Elisabeth
; APPLICANT: Jager, Elke
; APPLICANT: Knuth, Alexander
; APPLICANT: Old, Lloyd J.
; TITLE OF INVENTION: Isolated Nucleic Acid Molecules Encoding Cancer Associated Antigen
; TITLE OF INVENTION: Antigens Per Se, And Uses Thereof
; FILE REFERENCE: LUD 5538.1
; CURRENT APPLICATION NUMBER: US/09/270,437D
; CURRENT FILING DATE: 1999-03-16
; PRIOR APPLICATION NUMBER: 09/061,709
; PRIOR FILING DATE: 1998-04-17
; NUMBER OF SEQ ID NOS: 23
; SEQ ID NO 9
; LENGTH: 21
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
US-09-270-437D-9

Query Match 100.0%; Score 21; DB 5; Length 21;
Best Local Similarity 100.0%; Pred. No. 0.6;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAAAGTATCTTCAAGGAGGCC 21
DB 1 GAAAGTATCTTCAAGGAGGCC 21

RESULT 2
US-09-724-676-44306
; Sequence 44306, Application US/09724676
; GENERAL INFORMATION:
; APPLICANT: CompuGen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 CompuGen
; CURRENT APPLICATION NUMBER: US/09/724,676
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222

Mon Apr 21 10:30:57 2003

us-09-270-437d-9.rnpn

```

; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 44306
; LENGTH: 861
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-724-676-44306

Query Match      100.0%; Score 21; DB 6; Length 861;
Best Local Similarity 100.0%; Pred. No. 0.92;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAAAGTATCTTCAAGGACGCC 21
   |||||
Db 321 GAAAGTATCTTCAAGGACGCC 341

RESULT 3
US-09-724-676A-44306
; Sequence 44306, Application US/09724676A
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676A
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 44306
; LENGTH: 861
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-724-676A-44306

Query Match      100.0%; Score 21; DB 6; Length 861;
Best Local Similarity 100.0%; Pred. No. 0.92;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAAAGTATCTTCAAGGACGCC 21
   |||||
Db 321 GAAAGTATCTTCAAGGACGCC 341

RESULT 4
US-10-170-235-4888
; Sequence 4888, Application US/10170235
; GENERAL INFORMATION:
; APPLICANT: VENTIER, J. Craig
; TITLE OF INVENTION: KITS, SUCH AS NUCLEIC ACID ARRAYS, COMPRISING A MAJORITY OF HUMAN
; TITLE OF INVENTION: TRANSCRIPTS, FOR DETECTING EXPRESSION AND OTHER USES THEREOF
; FILE REFERENCE: CL001380
; CURRENT APPLICATION NUMBER: US/10/170,235
; CURRENT FILING DATE: 2003-03-17
; NUMBER OF SEQ ID NOS: 42514
; SEQ ID NO 4888
; LENGTH: 1558
; TYPE: DNA
; ORGANISM: HUMAN
US-10-170-235-4888

Query Match      100.0%; Score 21; DB 8; Length 1558;
Best Local Similarity 100.0%; Pred. No. 0.99;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAAAGTATCTTCAAGGACGCC 21
   |||||
Db 61 GAAAGTATCTTCAAGGACGCC 81

RESULT 5
US-10-313-986-347
; Sequence 347, Application US/10313986
; GENERAL INFORMATION:
; APPLICANT: Foy, Teresa M.

```

```

; APPLICANT: McNabb, Andria
; APPLICANT: Watanabe, Yoshihiro
; APPLICANT: Reed, Steven G.
; APPLICANT: Wang, Tongtong
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.455C19
; CURRENT APPLICATION NUMBER: US/10/313,986
; CURRENT FILING DATE: 2002-12-04
; NUMBER OF SEQ ID NOS: 560
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 347
; LENGTH: 1740
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-313-986-347

Query Match      100.0%; Score 21; DB 9; Length 1740;
Best Local Similarity 100.0%; Pred. No. 1;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAAAGTATCTTCAAGGACGCC 21
   |||||
Db 55 GAAAGTATCTTCAAGGACGCC 75

RESULT 6
US-10-313-986-478
; Sequence 478, Application US/10313986
; GENERAL INFORMATION:
; APPLICANT: Foy, Teresa M.
; APPLICANT: McNabb, Andria
; APPLICANT: Watanabe, Yoshihiro
; APPLICANT: Reed, Steven G.
; APPLICANT: Wang, Tongtong
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.455C19
; CURRENT APPLICATION NUMBER: US/10/313,986
; CURRENT FILING DATE: 2002-12-04
; NUMBER OF SEQ ID NOS: 560
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 478
; LENGTH: 1740
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-313-986-478

Query Match      100.0%; Score 21; DB 9; Length 1740;
Best Local Similarity 100.0%; Pred. No. 1;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAAAGTATCTTCAAGGACGCC 21
   |||||
Db 55 GAAAGTATCTTCAAGGACGCC 75

RESULT 7
US-10-313-986-483
; Sequence 483, Application US/10313986
; GENERAL INFORMATION:
; APPLICANT: Foy, Teresa M.
; APPLICANT: McNabb, Andria
; APPLICANT: Watanabe, Yoshihiro
; APPLICANT: Reed, Steven G.
; APPLICANT: Wang, Tongtong
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.455C19
; CURRENT APPLICATION NUMBER: US/10/313,986
; CURRENT FILING DATE: 2002-12-04
; NUMBER OF SEQ ID NOS: 560
; SOFTWARE: FastSEQ for Windows Version 4.0

```

; SEQ ID NO 483
; LENGTH: 1740
; TYPE: DNA
; ORGANISM: primate
US-10-313-986-483

Query Match 100.0%; Score 21; DB 9; Length 1740;
Best Local Similarity 100.0%; Pred. No. 1;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GAAAGTATCTTCAAGGACGCC 21
|||||
Db 55 GAAAGTATCTTCAAGGACGCC 75

RESULT 8

US-10-313-986-447
; Sequence 447, Application US/10313986
; GENERAL INFORMATION:
; APPLICANT: Foy, Teresa M.
; APPLICANT: McNabb, Andria
; APPLICANT: Watanabe, Yoshihiro
; APPLICANT: Reed, Steven G.
; APPLICANT: Wang, Tongtong
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.455C19
; CURRENT APPLICATION NUMBER: US/10/313,986
; CURRENT FILING DATE: 2002-12-04
; NUMBER OF SEQ ID NOS: 560
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 447
; LENGTH: 1743
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-313-986-447

Query Match 100.0%; Score 21; DB 9; Length 1743;
Best Local Similarity 100.0%; Pred. No. 1;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GAAAGTATCTTCAAGGACGCC 21
|||||
Db 55 GAAAGTATCTTCAAGGACGCC 75

RESULT 9

US-10-313-986-450
; Sequence 450, Application US/10313986
; GENERAL INFORMATION:
; APPLICANT: Foy, Teresa M.
; APPLICANT: McNabb, Andria
; APPLICANT: Watanabe, Yoshihiro
; APPLICANT: Reed, Steven G.
; APPLICANT: Wang, Tongtong
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.455C19
; CURRENT APPLICATION NUMBER: US/10/313,986
; CURRENT FILING DATE: 2002-12-04
; NUMBER OF SEQ ID NOS: 560
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 450
; LENGTH: 1743
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-313-986-450

Query Match 100.0%; Score 21; DB 9; Length 1743;
Best Local Similarity 100.0%; Pred. No. 1;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GAAAGTATCTTCAAGGACGCC 21

Db 55 GAAAGTATCTTCAAGGACGCC 75
|||||

RESULT 10

US-10-313-986-428
; Sequence 428, Application US/10313986
; GENERAL INFORMATION:
; APPLICANT: Foy, Teresa M.
; APPLICANT: McNabb, Andria
; APPLICANT: Watanabe, Yoshihiro
; APPLICANT: Reed, Steven G.
; APPLICANT: Wang, Tongtong
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.455C19
; CURRENT APPLICATION NUMBER: US/10/313,986
; CURRENT FILING DATE: 2002-12-04
; NUMBER OF SEQ ID NOS: 560
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 428
; LENGTH: 1764
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-313-986-428

Query Match 100.0%; Score 21; DB 9; Length 1764;
Best Local Similarity 100.0%; Pred. No. 1;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GAAAGTATCTTCAAGGACGCC 21
|||||
Db 76 GAAAGTATCTTCAAGGACGCC 96

RESULT 11

US-10-313-986-485
; Sequence 485, Application US/10313986
; GENERAL INFORMATION:
; APPLICANT: Foy, Teresa M.
; APPLICANT: McNabb, Andria
; APPLICANT: Watanabe, Yoshihiro
; APPLICANT: Reed, Steven G.
; APPLICANT: Wang, Tongtong
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.455C19
; CURRENT APPLICATION NUMBER: US/10/313,986
; CURRENT FILING DATE: 2002-12-04
; NUMBER OF SEQ ID NOS: 560
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 485
; LENGTH: 1799
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-313-986-485

Query Match 100.0%; Score 21; DB 9; Length 1799;
Best Local Similarity 100.0%; Pred. No. 1;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GAAAGTATCTTCAAGGACGCC 21
|||||
Db 74 GAAAGTATCTTCAAGGACGCC 94

RESULT 12

US-09-724-676-44307
; Sequence 44307, Application US/09724676
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen

us-09-270-437d-9.rnpn

Mon Apr 21 10:30:57 2003

; Sequence 44305, Application US/09724676A
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676A
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: Patentin version 3.2
; SEQ ID NO 44305
; LENGTH: 4060
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-724-676A-44305

Query Match 100.0%; Score 21; DB 6; Length 4060;
Best Local Similarity 100.0%; Pred. No. 1.1;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAAAGTATCTTCAAGGAGCGCC 21
|||
DB 321 GAAAGTATCTTCAAGGAGCGCC 341
|||

Search completed: April 18, 2003, 10:03:36
Job time : 238.023 secs

; CURRENT APPLICATION NUMBER: US/09/724,676
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: Patentin version 3.2
; SEQ ID NO 44307
; LENGTH: 2137
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-724-676-44307

Query Match 100.0%; Score 21; DB 6; Length 2137;
Best Local Similarity 100.0%; Pred. No. 1;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAAAGTATCTTCAAGGAGCGCC 21
|||
DB 321 GAAAGTATCTTCAAGGAGCGCC 341
|||

RESULT 13
US-09-724-676A-44307
; Sequence 44307, Application US/09724676A
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676A
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: Patentin version 3.2
; SEQ ID NO 44307
; LENGTH: 2137
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-724-676A-44307

Query Match 100.0%; Score 21; DB 6; Length 2137;
Best Local Similarity 100.0%; Pred. No. 1;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAAAGTATCTTCAAGGAGCGCC 21
|||
DB 321 GAAAGTATCTTCAAGGAGCGCC 341
|||

RESULT 14
US-09-724-676-44305
; Sequence 44305, Application US/09724676
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: Patentin version 3.2
; SEQ ID NO 44305
; LENGTH: 4060
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-724-676-44305

Query Match 100.0%; Score 21; DB 6; Length 4060;
Best Local Similarity 100.0%; Pred. No. 1.1;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAAAGTATCTTCAAGGAGCGCC 21
|||
DB 321 GAAAGTATCTTCAAGGAGCGCC 341
|||

RESULT 15
US-09-724-676A-44305

GenCore version 5.1.4 p5 4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 18, 2003, 04:47:40 ; Search time 731.182 Seconds
(without alignments)
465.145 Million cell updates/sec

Title: US-09-270-437D-9

Perfect score: 21
Sequence: 1 gaagatatttcaggagcgc 21

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST:*

- 1: em_estba:*
- 2: em_esthum:*
- 3: em_estin:*
- 4: em_estm:*
- 5: em_estov:*
- 6: em_estpl:*
- 7: em_estro:*
- 8: em_htc:*
- 9: gb_est1:*
- 10: gb_est2:*
- 11: gb_htc:*
- 12: gb_est3:*
- 13: gb_est4:*
- 14: gb_est5:*
- 15: em_estfun:*
- 16: em_estom:*
- 17: gb_gss:*
- 18: em_gss_hum:*
- 19: em_gss_inv:*
- 20: em_gss_pln:*
- 21: em_gss_vrt:*
- 22: em_gss_fun:*
- 23: em_gss_mam:*
- 24: em_gss_mus:*
- 25: em_gss_other:*
- 26: em_gss_pro:*
- 27: em_gss_rod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	21	100.0	360	10	BE560346
2	21	100.0	361	10	BE269531
3	21	100.0	361	10	BE560268
4	21	100.0	366	10	BE396918
5	21	100.0	371	10	BE560799
6	21	100.0	384	10	BE396917

7	21	100.0	477	9	AI682259
8	21	100.0	776	10	BE513300
9	21	100.0	910	12	BE881835
10	21	100.0	918	14	BQ437571
11	21	100.0	3100	11	BC019258
12	20	95.2	121	9	AA836048
13	18.4	87.6	178	9	AA836048
14	18.4	87.6	347	9	AA836048
15	18.4	87.6	356	14	AA003676
16	18.4	87.6	446	10	W11682
17	18.4	87.6	489	9	AA318369
18	18.4	87.6	573	10	AA547497
19	18.4	87.6	600	12	BB639276
20	18.4	87.6	600	13	BB639276
21	18.4	87.6	601	9	BB639276
22	18.4	87.6	601	10	BB639276
23	18.4	87.6	632	10	BB612512
24	18.4	87.6	648	10	BB657710
25	18.4	87.6	712	11	BB657710
26	18.4	87.6	749	14	BQ180708
27	18.4	87.6	895	14	BQ30512
28	18.4	87.6	930	14	BQ895257
29	18.4	87.6	2202	11	BQ895257
30	17.8	84.8	195	17	AK011689
31	17.8	84.8	325	13	AK011689
32	17.8	84.8	441	12	AK011689
33	17.4	82.9	195	9	BF776116
34	17.4	82.9	499	10	BF776116
35	17	81.0	332	17	AT000448
36	17	81.0	334	13	BB757949
37	17	81.0	418	10	AZ743062
38	17	81.0	612	17	BJ158403
39	17	81.0	631	13	AW497094
40	17	81.0	718	17	AZ667042
41	17	81.0	763	17	BJ165180
42	17	81.0	781	17	BH377427
43	16.8	80.0	108	12	BH527295
44	16.8	80.0	316	12	BOHEM88TF
45	16.8	80.0	322	12	BOH0833TF

ALIGNMENTS

RESULT 1
BE560346
LOCUS BE560346
DEFINITION 601346615F1 NIH_MGC_8 Homo sapiens linear EST 15-AUG-2000
ACCESSION BE560346
VERSION BE560346.1
KEYWORDS GI:9804066
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 360)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-rc@mail.nih.gov
Tissue Procurement: Louis M. Staudt, M.D., Ph.D.
cDNA Library Preparation: Ling Hong/Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
Plate: L1CM379 row: m column: 03
High quality sequence start: 4
High quality sequence stop: 360.
Location/Qualifiers 1..360

AI682259 wa71105.X
BE513300 601315375
BE881835 601504962
BQ437571 AGENCOURT
BC019258 Homo sapi
AA836048 Qd43C04.8
AA836048 vm64H08.8
AA003676 mg59C10.X
W11682 me34C05.X1
AA318369 um95E08.Y
AA547497 vk71G05.8
BB639276 BB639276
BB639276 0221-68 M
B1986515 3172-67 M
AA522010 vh78b12.X
BB611741 BB611741
BB612512 BB612512
BB657710 BB657710
AK011797 Mus muscu
BQ180708 UI-M-EXO-
BQ30512 AGENCOURT
BQ895257 AGENCOURT
AK011689 Mus muscu
AQ347018 RPI11-11
B1477406 da14H07.
BF776116 286477 MA
AT000448 AT000448
BB757949 BB757949
AZ743062 RPCI-24-9
BJ158403 BJ158403
AW497094 9A53G05.Y
AZ667042 IM0549N03
BJ165180 BJ165180
BH377427 AG-ND-135
BH527295 BOHEM88TF
BH432699 BOH0833TF
BF377514 MR4-TN011
BG233936 Gaa3a01.
BG373492 UI-R-CSO-

us-09-270-437d-9.rst

Mon Apr 21 10:30:58 2003

Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GAAAGTATCTTCAAGGAGCC 21
 |||||
 Db 286 GAAAGTATCTTCAAGGAGCC 306

RESULT 3
 BE560268 361 bp mRNA linear EST 15-AUG-2000
 LOCUS 601346086F1 NIH_MGC_8 Homo sapiens cDNA clone IMAGE:3678885 5',
 DEFINITION mRNA sequence.
 ACCESSION BE560268
 VERSION BE560268.1 GI:9803988
 KEYWORDS EST.
 SOURCE human.

ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 361)
 AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: Louis M. Staudt, M.D., Ph.D.
 cDNA Library Preparation: Ling Hong/Rubin Laboratory
 DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
 Plate: LLCM356 row: 1 column: 22
 High quality sequence start: 4
 High quality sequence stop: 361.

FEATURES
 source
 1..361
 Location/Qualifiers
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:3678885"
 /clone_lib="NIH_MGC_8"
 /tissue_type="Burkitt lymphoma"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: lymph; Vector: pOTB7; Site 1: XhoI; Site 2:
 EcoRI; cDNA made by oligo-dt priming. Directionally
 cloned into EcoRI/XhoI sites using the following 5'
 adaptor: GGCACGAG(G). Size-selected >500bp for average
 insert size 1.8kb. Library constructed by Ling Hong in
 the laboratory of Gerald M. Rubin (University of
 California, Berkeley) using ZAP-cDNA synthesis kit
 (Stratagene) and Superscript II RT (Life Technologies)."

BASE COUNT 88 a 85 c 75 g 113 t
 ORIGIN
 Query Match 100.0%; Score 21; DB 10; Length 361;
 Best Local Similarity 100.0%; Pred. No. 3.4;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GAAAGTATCTTCAAGGAGCC 21
 |||||
 Db 286 GAAAGTATCTTCAAGGAGCC 306

RESULT 4
 BE396918 366 bp mRNA linear EST 21-JUL-2000
 LOCUS 601289477F1 NIH_MGC_8 Homo sapiens cDNA clone IMAGE:3620058 5',
 DEFINITION mRNA sequence.
 ACCESSION BE396918
 VERSION BE396918.1 GI:9342283
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Query Match 100.0%; Score 21; DB 10; Length 361;
 Best Local Similarity 100.0%; Pred. No. 3.4;

/organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:3678885"
 /clone_lib="NIH_MGC_8"
 /tissue_type="Burkitt lymphoma"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: lymph; Vector: pOTB7; Site 1: XhoI; Site 2:
 EcoRI; cDNA made by oligo-dt priming. Directionally
 cloned into EcoRI/XhoI sites using the following 5'
 adaptor: GGCACGAG(G). Size-selected >500bp for average
 insert size 1.8kb. Library constructed by Ling Hong in
 the laboratory of Gerald M. Rubin (University of
 California, Berkeley) using ZAP-cDNA synthesis kit
 (Stratagene) and Superscript II RT (Life Technologies)."

BASE COUNT 88 a 85 c 75 g 112 t
 ORIGIN
 Query Match 100.0%; Score 21; DB 10; Length 360;
 Best Local Similarity 100.0%; Pred. No. 3.4;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GAAAGTATCTTCAAGGAGCC 21
 |||||
 Db 286 GAAAGTATCTTCAAGGAGCC 306

RESULT 2
 BE269531 361 bp mRNA linear EST 13-JUL-2000
 LOCUS 601184745F1 NIH_MGC_8 Homo sapiens cDNA clone IMAGE:3542483 5',
 DEFINITION mRNA sequence.

ACCESSION BE269531
 VERSION BE269531.1 GI:9143153
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 361)
 AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: Louis M. Staudt, M.D., Ph.D.
 cDNA Library Preparation: Ling Hong/Rubin Laboratory
 DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
 Plate: LLCM234 row: 1 column: 12
 High quality sequence stop: 361.

FEATURES
 source
 1..361
 Location/Qualifiers
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:3542483"
 /clone_lib="NIH_MGC_8"
 /tissue_type="Burkitt lymphoma"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: lymph; Vector: pOTB7; Site 1: XhoI; Site 2:
 EcoRI; cDNA made by oligo-dt priming. Directionally
 cloned into EcoRI/XhoI sites using the following 5'
 adaptor: GGCACGAG(G). Size-selected >500bp for average
 insert size 1.8kb. Library constructed by Ling Hong in
 the laboratory of Gerald M. Rubin (University of
 California, Berkeley) using ZAP-cDNA synthesis kit
 (Stratagene) and Superscript II RT (Life Technologies)."

BASE COUNT 88 a 85 c 75 g 113 t
 ORIGIN
 Query Match 100.0%; Score 21; DB 10; Length 361;
 Best Local Similarity 100.0%; Pred. No. 3.4;


```

REFERENCE
AUTHORS NIH-MGC http://mgc.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Louis M. Staudt, M.D., Ph.D.
CDNA Library Preparation: Ling Hong/Rubin Laboratory
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
Plate: LLCW291 row: 1 column: 19
High quality sequence start: 26
High quality sequence stop: 366.
FEATURES
source
1..366
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:3620058"
/clone_lib="NIH_MGC_8"
/tissue_type="Burkitt lymphoma"
/lab_host="DH10B (phage-resistant)"
/notes="Organ: lymph; Vector: pOTB7; Site: 1: XhoI; Site 2:
EcoRI; CDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adaptor: GGCACGAG(G). Size-selected >500bp for average
insert size 1.8kb. Library constructed by Ling Hong in
the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)."
BASE COUNT 89 a 87 c 76 g 114 t
ORIGIN
Query Match 100.0%; Score 21; DB 10; Length 366;
Best Local Similarity 100.0%; Pred. No. 3.5;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GAAAGTATCTTCAAGGAGCC 21
|||||
Db 291 GAAAGTATCTTCAAGGAGCC 311

RESULT 5
BE560799 371 bp mRNA linear EST 15-AUG-2000
LOCUS 601346251F1 NIH_MGC_8 Homo sapiens CDNA clone IMAGE:3679476 5',
mRNA sequence.
ACCESSION BE560799.1 GI:9804519
VERSION BE560799.1
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 371)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Louis M. Staudt, M.D., Ph.D.
CDNA Library Preparation: Ling Hong/Rubin Laboratory
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
Plate: LLCW358 row: e column: 13
High quality sequence start: 6
High quality sequence stop: 371.
FEATURES
source
1..371
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:3679476"
/clone_lib="NIH_MGC_8"
/tissue_type="Burkitt lymphoma"
/lab_host="DH10B (phage-resistant)"
/notes="Organ: lymph; Vector: pOTB7; Site 1: XhoI; Site 2:
EcoRI; CDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adaptor: GGCACGAG(G). Size-selected >500bp for average
insert size 1.8kb. Library constructed by Ling Hong in
the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)."
BASE COUNT 90 a 93 c 80 g 121 t
ORIGIN
Query Match 100.0%; Score 21; DB 10; Length 384;

```

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:3679476"
/clone_lib="NIH_MGC_8"
/tissue_type="Burkitt lymphoma"
/lab_host="DH10B (phage-resistant)"
/notes="Organ: lymph; Vector: pOTB7; Site: 1: XhoI; Site 2:
EcoRI; CDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adaptor: GGCACGAG(G). Size-selected >500bp for average
insert size 1.8kb. Library constructed by Ling Hong in
the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)."
BASE COUNT 90 a 94 c 75 g 112 t
ORIGIN
Query Match 100.0%; Score 21; DB 10; Length 371;
Best Local Similarity 100.0%; Pred. No. 3.5;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GAAAGTATCTTCAAGGAGCC 21
|||||
Db 296 GAAAGTATCTTCAAGGAGCC 316

RESULT 6
BE396917 384 bp mRNA linear EST 21-JUL-2000
LOCUS 601289473F1 NIH_MGC_8 Homo sapiens CDNA clone IMAGE:3619866 5',
mRNA sequence.
ACCESSION BE396917
VERSION BE396917.1 GI:9342282
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 384)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Louis M. Staudt, M.D., Ph.D.
CDNA Library Preparation: Ling Hong/Rubin Laboratory
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
Plate: LLCM291 row: a column: 19
High quality sequence start: 49
High quality sequence stop: 384.
FEATURES
source
1..384
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:3619866"
/clone_lib="NIH_MGC_8"
/tissue_type="Burkitt lymphoma"
/lab_host="DH10B (phage-resistant)"
/notes="Organ: lymph; Vector: pOTB7; Site 1: XhoI; Site 2:
EcoRI; CDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adaptor: GGCACGAG(G). Size-selected >500bp for average
insert size 1.8kb. Library constructed by Ling Hong in
the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)."
BASE COUNT 90 a 93 c 80 g 121 t
ORIGIN
Query Match 100.0%; Score 21; DB 10; Length 384;

```

ORGANISM	Homo sapiens	
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	
AUTHORS	NIH-MGC http://mgc.nci.nih.gov/ .	
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)	
JOURNAL	Unpublished (1999)	
COMMENT	Contact: Robert Strausberg, Ph.D. Email: cgapbs-remail.nih.gov Tissue Procurement: Louis M. Staudt, M.D., Ph.D. cDNA Library Preparation: Ling Hong/Rubin Laboratory cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Incyte Genomics, Inc. Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov Plate: LLCM327 row: f column: 19 High quality sequence stop: 615.	
FEATURES	Location/Qualifiers	
Source	1..776	
	/organism="Homo sapiens"	
	/db_xref="taxon:9606"	
	/clone="IMAGE:3633810"	
	/clone_lib="NIH_MGC 8"	
	/tissue_type="Burkitt lymphoma"	
	/lab_host="DH10B (phage-resistant)"	
	/note="Organ: lymph; Vector: pOTB7; Site: 1: XhoI; Site: 2: EcoRI; cDNA made by oligo-dt priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies)."	
BASE COUNT	176 a	178 c 228 g 194 t
ORIGIN		
Query Match	100.0%;	Score 21; DB 10; Length 776;
Best Local Similarity	100.0%;	Pred. No. 4.9;
Matches	21; Conservative	0; Mismatches 0; Indels 0; Gaps 0;
QY	1	GAAAGTATCTTCAAGGAGCC 21
Db	290	GAAAGTATCTTCAAGGAGCC 310
RESULT 9		
BE881835		
LOCUS	BE881835 910 bp mRNA linear EST 20-OCT-2000	
DEFINITION	G0150492F1 NIH_MGC_71 Homo sapiens cDNA clone IMAGE:3906590 5', mRNA sequence.	
ACCESSION	BE881835	
VERSION	BE881835.1 GI:10330611	
KEYWORDS	EST.	
SOURCE	human.	
ORGANISM	Homo sapiens	
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	
AUTHORS	NIH-MGC http://mgc.nci.nih.gov/ .	
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)	
JOURNAL	Unpublished (1999)	
COMMENT	Contact: Robert Strausberg, Ph.D. Email: cgapbs-remail.nih.gov Tissue Procurement: ATCC cDNA Library Preparation: Life Technologies, Inc. cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Incyte Genomics, Inc. Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov Plate: L1AM9715 row: 1 column: 15 High quality sequence start: 56 High quality sequence stop: 746.	

```

Best Local Similarity 100.0%; Pred. No. 3.5;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAAAGTATCTTCAAGGACGCC 21
|||||
DB 309 GAAAGTATCTTCAAGGACGCC 329
|||||

RESULT 7
AI682259 477 bp mRNA linear EST 07-MAR-2000
wa71h05.x1 Soares NFL T GBC S1 Homo sapiens cDNA clone IMAGE:2301657 3' similar to TR:00425 000425 PUTATIVE RNA BINDING PROTEIN KOC :, mRNA sequence.
AI682259 1 GI:4892441
EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 477)
NCI-CCGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
This clone is available royalty-free through LNL; contact the IMAGE Consortium (info@image.lnl.gov) for further information.
Insert length: 608 Std Error: 0.00
Seq primer: -40UP from Gibco
High quality sequence stop: 427.
Location/Qualifiers
1. 477
/organism="Homo sapiens"
/db xref="taxon:9606"
/clone="IMAGE:2301657"
/clone_lib="Soares_NFL_T_GBC_S1"
/lab_host="DH10B"
/notes="Organ: pooled; Vector: p773D-Pac (Pharmacia) with a modified polylinker; Site_1: Not 1; Site_2: Eco RI; Equal amounts of plasmid DNA from three normalized libraries (fetal lung MbHL19w, testis NHT, and B-cell NCI CGAP GCBI) were mixed, and ss circles were made in vitro. Following HAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from pools of 5,000 clones made from the same 3 libraries. The pools consisted of I.M.A.G.E. clones 297480-302087, 682632-687239, 726408-728711, and 729096-731399. Subtraction by Bento Soares and M. Fatima Bonaldo."
BASE COUNT 122 a 99 c 102 g 153 t 1 others
ORIGIN

Query Match 100.0%; Score 21; DB 9; Length 477;
Best Local Similarity 100.0%; Pred. No. 3.9;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAAAGTATCTTCAAGGACGCC 21
|||||
DB 355 GAAAGTATCTTCAAGGACGCC 375
|||||

RESULT 8
BE513300 776 bp mRNA linear EST 07-AUG-2000
LOCUS BE513300
DEFINITION 601315375F1 NIH_MGC_8 Homo sapiens cDNA clone IMAGE:3633810 5', mRNA sequence.
ACCESSION BE513300
VERSION BE513300.1 GI:9720512
KEYWORDS EST.
SOURCE human.

```

```

FEATURES
Source
Location/Qualifiers
1..910
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:3908590"
/clone_lib="NIH_MGC_71"
/tissue_type="leiomyosarcoma"
/lab_host="DH10B (phage-resistant)"
/notes="Organ: uterus; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 2.1 kb."
BASE COUNT      252 a      222 c      241 g      194 t      1 others
ORIGIN
Query Match      100.0%; Score 21; DB 12; Length 910;
Best Local Similarity 100.0%; Pred. No. 5.3;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAAAGTATCTTCAAGGACGCC 21
|||||
Db 311 GAAAGTATCTTCAAGGACGCC 331

RESULT 10
LOCUS      BQ437571
DEFINITION AGENCOURT_7898330 NIH_MGC_67 Homo sapiens cDNA clone IMAGE:6154021
5', mRNA sequence.
ACCESSION  BQ437571
VERSION     BQ437571.1 GI:21176647
KEYWORDS    EST.
SOURCE      human.
ORGANISM    Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 918)
NIH-MGC http://mgc.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1AM13494 row: 9 column: 14
High quality sequence stop: 555.
FEATURES
Source
Location/Qualifiers
1..918
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:6154021"
/clone_lib="NIH_MGC_67"
/tissue_type="rctingblastoma"
/lab_host="DH10B (phage-resistant)"
/notes="Organ: eye; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 1.75 kb. Library constructed by Life
Technologies."
BASE COUNT      237 a      225 c      232 g      224 t
ORIGIN
Query Match      100.0%; Score 21; DB 14; Length 918;
Best Local Similarity 100.0%; Pred. No. 5.3;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAAAGTATCTTCAAGGACGCC 21
|||||
Db 322 GAAAGTATCTTCAAGGACGCC 342

RESULT 11
LOCUS      BC019258
DEFINITION Homo sapiens, Similar to IGF-II mRNA-binding protein 3, clone
IMAGE:3542483, mRNA.
ACCESSION  BC019258
VERSION     BC019258.1 GI:17939534
KEYWORDS    HTC.
SOURCE      Homo sapiens.
ORGANISM    Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 3100)
Strausberg.R.
Direct Submission.
Submitted (13-DEC-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA.
NIH-MGC Project URL: http://mgc.nci.nih.gov
Contact: MGC help desk
Email: cgabbs-remail.nih.gov
Tissue Procurement: Louis M. Staudt, M.D., Ph.D.
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Genome Sequence Centre,
BC Cancer Agency, Vancouver, BC, Canada
info@bcgsc.bc.ca
Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield,
Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin,
Leticia Heiao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Soo
Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven
Ness, Fawan Pandoh, Anna-Liisa Prabhu, Parvaneh Saeedi, Jacqueline
Schein, Duane Smailus, Michael Smith, Lorraine Spence, Jeff Stott,
Michael Thorne, Miranada Teai, Natasja van den Bosch, Jill Vardy,
George Yang, Scott Zuyderduyn, Marco Marra.
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAL Plate: 8 Row: c Column: 6
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 5729900
This clone has the following problem: incomplete processing.
FEATURES
Source
Location/Qualifiers
1..3100
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:3542483"
/tissue_type="Lymph, Burkitt lymphoma"
/clone_lib="NIH_MGC_8"
/lab_host="DH10B-R"
/notes="Vector: pOTB7"
BASE COUNT      748 a      674 c      873 g      805 t
ORIGIN
Query Match      100.0%; Score 21; DB 11; Length 3100;
Best Local Similarity 100.0%; Pred. No. 9.4;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAAAGTATCTTCAAGGACGCC 21
|||||
Db 305 GAAAGTATCTTCAAGGACGCC 325

RESULT 12
LOCUS      AA836048
DEFINITION AA836048 121 bp mRNA linear EST 31-MAR-1998
OD43C04.s1 NCI CGAP GCB1 Homo sapiens cDNA clone IMAGE:1370694
similar to TR:000425 O00425 PUTATIVE RNA BINDING PROTEIN KOC ;,
mRNA sequence.
ACCESSION  AA836048
VERSION     AA836048.1 GI:2910367

```

KEYWORDS EST.
SOURCE Homo sapiens
ORGANISM Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 121)
AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgaps@mail.nih.gov
 Tissue Procurement: Louis M. Staudt, M.D., Ph.D., David Allman, Ph.D., Gerald Marti, M.D.
 CDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima Bonaldo, Ph.D.
 CDNA Library Arrayed by: Greg Lennon, Ph.D.
 DNA Sequencing by: Washington University Genome Sequencing Center
 Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: www-bio.llnl.gov/bbrp/image/image.html
 Insert Length: 730 Std Error: 0.00
 Seq primer: -40ml3 fwd. ET from Amersham
 High quality sequence stop: 118.
FEATURES Location/Qualifiers
 1..121
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:1370694"
 /clone_lib="NCI CGAP GCB1"
 /tissue_type="germinal center B cell"
 /lab_host="DH10B"
 /note="Vector: p7T3D-Pac (Pharmacia) with a modified polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA was prepared from human tonsillar cells enriched for germinal center B cells by flow sorting (CD20+, IgD-), provided by Dr. Louis M. Staudt (NCI), Dr. David Allman (NCI) and Dr. Gerald Marti (CBER). cDNA synthesis was primed with a Not I - oligo(dT) primer [5'-TGTTACCAATCTGAAGTGGGCGGCGCTCAATTTTCTTTT-3'}. Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified p7T3 vector. Library went through one round of normalization, and was constructed by Bento Soares and M. Fatima Bonaldo."
 37 a 34 c 20 g 30 t
BASE COUNT
ORIGIN
 Query Match 95.2%; Score 20; DB 9; Length 121;
 Best Local Similarity 100.0%; Pred. No. 6.8;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2 AAAGTATCTTCAAGGACGCC 21
 |||||
Db 91 AAAGTATCTTCAAGGACGCC 110
RESULT 13
AA683695/c
LOCUS 178 bp mRNA linear EST 09-DEC-1997
DEFINITION vm64h08.s1 Knowles Solter mouse 2 cell Mus musculus cDNA clone IMAGE:1003071 5' similar to TR:000425 O00425 PUTATIVE RNA BINDING PROTEIN KOC ;, mRNA sequence.
ACCESSION AA683695
VERSION AA683695.1 GI:2671086
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
REFERENCE 1 (bases 1 to 178)
AUTHORS Marra, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T., Geisel, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, M., Schellenberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B., Theising, B., Wylie, T., Lennon, G., Soares, B., Wilson, R. and Waterston, R.
 The WashU-HMMI Mouse EST Project
 Unpublished (1996)
 Contact: Marra M/Mouse EST Project
 WashU-HMMI Mouse EST Project
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: mousseest@watson.wustl.edu
 This clone is available royalty-free through LLNL ; contact the

Schellenberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B., Theising, B., Wylie, T., Lennon, G., Soares, B., Wilson, R. and Waterston, R.
 The WashU-HMMI Mouse EST Project
 Unpublished (1996)
 Contact: Marra M/Mouse EST Project
 WashU-HMMI Mouse EST Project
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: mousseest@watson.wustl.edu
 This clone is available royalty-free through LLNL ; contact the
FEATURES Location/Qualifiers
 1..178
 /organism="Mus musculus"
 /strain="C57BL/6J x DBA/2J F1"
 /db_xref="taxon:10090"
 /clone="IMAGE:1003071"
 /clone_lib="Knowles Solter mouse 2 cell"
 /tissue_type="embryo"
 /dev_stage="2-cell"
 /lab_host="DH10B"
 /note="Organ: embryo; Vector: pBluescribe (modified); Site 1: MluI; Site 2: SalI; Cloned unidirectionally from mRNA prepared from 13,500 2-cell stage embryos. Primer: SalI (dT): 5'-CGTGCACGTCGACGCTTTT-3'. cDNAs were cloned into the MluI/SalI sites of a modified pBluescribe vector using commercial linkers (NEB).
 Average insert size: 1.2 Kb."
 43 a 45 c 52 g 38 t
BASE COUNT
ORIGIN
 Query Match 87.6%; Score 18.4; DB 9; Length 178;
 Best Local Similarity 95.0%; Pred. No. 55;
 Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 GAAAGTATCTTCAAGGACGC 20
 |||||
Db 60 GAAAGTCTCTTCAAGGACGC 41
RESULT 14
AA003676
LOCUS 347 bp mRNA linear EST 22-JUL-1996
DEFINITION ng59c10.r1 Soares mouse embryo NbME13.5 14.5 Mus musculus cDNA clone IMAGE:437298 5', mRNA sequence.
ACCESSION AA003676
VERSION AA003676.1 GI:1447248
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 347)
AUTHORS Marra, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T., Geisel, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, M., Schellenberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B., Theising, B., Wylie, T., Lennon, G., Soares, B., Wilson, R. and Waterston, R.
 The WashU-HMMI Mouse EST Project
 Unpublished (1996)
 Contact: Marra M/Mouse EST Project
 WashU-HMMI Mouse EST Project
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: mousseest@watson.wustl.edu
 This clone is available royalty-free through LLNL ; contact the

IMAGE Consortium (info@image.llnl.gov) for further information.
 MGI:262634

Seq primer: ETPPrimer
 High quality sequence stop: 333.

FEATURES

Location/Qualifiers
 1..347

/organism="Mus musculus"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="IMAGE:437298"
 /clone_lib="Soares mouse embryo NbME13.5 14.5"
 /sex="unknown"
 /tissue_type="embryo"
 /dev_stage="13.5-14.5dpc total fetus"
 /lab_hosts="DH10B"

/note="Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5', TGTTACCAATCTGAAGTGGAGCGCGCGGAATTTTTTTTTTTTTTTT T 3'], on equal amounts of mRNA from 2 13.5dpc and 2 14.5dpc embryos [total RNA provided by Minoru Ko, Wayne State Univ., from 2]; double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT7T3 vector. Library went through one round of normalization, and was constructed by Bento Soares and M.Fatima Bonaldo. "

BASE COUNT 81 a 92 c 100 g 74 t

ORIGIN

Query Match 87.6%; Score 18.4; DB 9; Length 347;
 Best Local Similarity 95.0%; Pred. No. 75;
 Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GAAAGTATCTTCAAGGACGC 20

Db 168 GAAAGTGTCTTCAAGGACGC 187

RESULT 15

W71682

LOCUS me34c05.r1 Soares mouse embryo NbME13.5 14.5 Mus musculus cDNA
 DEFINITION clone IMAGE:389384 5', mRNA sequence.

ACCESSION W71682

VERSION W71682.1 GI:1380423

KEYWORDS EST.

SOURCE house mouse.

ORGANISM

Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 356)
 Marra, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T.,
 Geisler, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, M.,
 Schellenberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B.,
 Theising, B., Wylie, T., Lennon, G., Soares, B., Wilson, R. and
 Waterston, R.

PROJECT The WashU-HMI Mouse EST Project

UNPUBLISHED (1996)

CONTACT: Marra M/Mouse EST Project

WashU-HMI Mouse EST Project

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108

Tel: 314 286 1800

Fax: 314 286 1810

Email: mouseest@watson.wustl.edu

This clone is available royalty-free through LLNL; contact the

IMAGE Consortium (info@image.llnl.gov) for further information.

MGI:241216

Seq primer: ETPPrimer

High quality sequence stop: 354.

Location/Qualifiers

1..356

FEATURES

Location/Qualifiers
 1..356

/organism="Mus musculus"

/strain="C57BL/6J"

/db_xref="taxon:10090"

/clone="IMAGE:389384"

/clone_lib="Soares mouse embryo NbME13.5 14.5"

/sex="unknown"

/tissue_type="embryo"

/dev_stage="13.5-14.5dpc total fetus"

/lab_hosts="DH10B"

/note="Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5', TGTTACCAATCTGAAGTGGAGCGCGCGGAATTTTTTTTTTTTTTTT T 3'], on equal amounts of mRNA from 2 13.5dpc and 2 14.5dpc embryos [total RNA provided by Minoru Ko, Wayne State Univ., from 2]; double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT7T3 vector. Library went through one round of normalization, and was constructed by Bento Soares and M.Fatima Bonaldo. "

BASE COUNT 79 a 93 c 93 g 91 t

ORIGIN

Query Match 87.6%; Score 18.4; DB 14; Length 356;
 Best Local Similarity 95.0%; Pred. No. 76;
 Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GAAAGTATCTTCAAGGACGC 20

Db 322 GAAAGTGTCTTCAAGGACGC 341

Search completed: April 18, 2003, 07:31:44
 Job time : 737.348 secs

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 18, 2003, 04:46:26 ; Search time 292 Seconds
(without alignments)
2192.677 Million cell updates/sec

Title: US-09-270-437D-10

Perfect score: 22
Sequence: 1 ctgcaagggttttctgtggcg 22

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl:

1: gb.ba.*

2: gb.htg.*

3: gb.in.*

4: gb.ov.*

5: gb.pat.*

6: gb.pl.*

7: gb.pr.*

8: gb.ro.*

9: gb.sts.*

10: gb.un.*

11: gb.vi.*

12: gb.vi.*

13: gb.vi.*

14: gb.vi.*

15: em.ba.*

16: em.fun.*

17: em.hum.*

18: em.in.*

19: em.mu.*

20: em.mu.*

21: em.or.*

22: em.ov.*

23: em.pat.*

24: em.ph.*

25: em.pl.*

26: em.ro.*

27: em.sts.*

28: em.un.*

29: em.vi.*

30: em.htg.hum.*

31: em.htg.inv.*

32: em.htg.other.*

33: em.htg.mus.*

34: em.htg.pln.*

35: em.htg.rod.*

36: em.htg.mam.*

37: em.htg.vrt.*

38: em.ey.*

39: em.htgo.hum.*

40: em.htgo.mus.*

41: em.htgo.other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
C 1	22	100.0	1740	6	AX365954	Sequence
C 2	22	100.0	1743	6	AX366054	Sequence
C 3	22	100.0	1743	6	AX366057	Sequence
C 4	22	100.0	1764	6	AX366035	Sequence
C 5	22	100.0	4155	9	HSU76705	U76705 Human putat
C 6	22	100.0	4159	9	ARI17863	ARI17863 Sequence
C 7	22	100.0	4171	9	AF117108	AF117108 Homo sapi
C 8	22	100.0	4181	6	AX333233	Sequence
C 9	22	100.0	4181	9	HSU7188	Sequence
C 10	22	100.0	4181	9	HSU7188	Sequence
C 11	22	100.0	4602	6	AX397963	U97188 Homo sapien
C 12	22	100.0	104668	9	AC021876	Sequence
C 13	22	100.0	158105	2	AC023375	Sequence
C 14	18.8	85.5	98951	2	AC023375	Homo sapi
C 15	18.8	85.5	200956	9	AC090069	Homo sapi
C 16	18.8	85.5	208197	9	AC104165	Homo sapi
C 17	18.4	83.6	111764	9	AC009107	Homo sapi
C 18	18.4	83.6	156422	2	AC106309	Homo sapi
C 19	18.4	83.6	159082	9	AC025947	Human DNA
C 20	18.4	83.6	163404	2	AL355308	Rattus no
C 21	18.4	83.6	172777	9	AC026395	Homo sapi
C 22	18	81.8	79554	8	NCB11B23	Homo sapi
C 23	17.8	80.9	32145	2	AC113360	Neurospor
C 24	17.8	80.9	58511	2	AC097402	Homo sapi
C 25	17.8	80.9	91200	9	AP001152	Rattus no
C 26	17.8	80.9	140666	9	AC120114	Homo sapi
C 27	17.8	80.9	141085	2	AC097250	Homo sapi
C 28	17.8	80.9	147556	2	AC011007	Rattus no
C 29	17.8	80.9	157481	9	AC093512	Homo sapi
C 30	17.8	80.9	173235	2	AC023756	Homo sapi
C 31	17.8	80.9	183666	9	AC009487	Homo sapi
C 32	17.8	80.9	189947	2	AC02035	Homo sapi
C 33	17.8	80.9	191076	2	AC115786	Homo sapi
C 34	17.8	80.9	203770	9	AC082651	Mus muscu
C 35	17.8	80.9	235669	2	AC096932	Homo sapi
C 36	17.4	79.1	56509	9	AL353700	Rattus no
C 37	17.4	79.1	83596	2	AC095682	Homo sapi
C 38	17.4	79.1	94180	2	AC106942	Rattus no
C 39	17.4	79.1	150722	2	AC120491	Rattus no
C 40	17.4	79.1	154218	9	AC020663	Rattus no
C 41	17.4	79.1	165777	2	AC111441	Homo sapi
C 42	17.4	79.1	168085	9	AC027687	Homo sapi
C 43	17.4	79.1	181347	2	AC117303	Rattus no
C 44	17.4	79.1	214042	2	AC093028	Homo sapi
C 45	17.4	79.1	258120	2	AC114408	Mus muscu

ALIGNMENTS

RESULT 1
AX365954/c
LOCUS AX365954
DEFINITION Sequence 347 from Patent WO0200174.
ACCESSION AX365954
VERSION AX365954.1 GI:18697455
KEYWORDS human.
SOURCE Homo sapiens
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1
Wang, T., Wang, A., Skeiky, Y.A., Li, S.X., Kalos, M.D., Henderson, R.A.,
McNeill, P.D., Fanger, N., Retter, M.W., Marnierakis, M., Fanger, G.R.,
Vedwick, T.S., Carter, D., Watanabe, Y. and Peckham, D.W.

Mon Apr 21 10:29:39 2003

Compositions and methods for the therapy and diagnosis of lung

TITLE Cancer
JOURNAL Patent: WO 0200174-A 347 03-JAN-2002;
FEATURES CORIXA CORPORATION (US)
 source Location/Qualifiers
 1. .1743
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
BASE COUNT 527 a 406 c 418 g 392 t
ORIGIN
 Query Match 100.0%; Score 22; DB 6; Length 1743;
 Best Local Similarity 100.0%; Pred. No. 5.6;
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 CTGCAAGGGGTTTGTCTGGCG 22
 Db 498 CTGCAAGGGGTTTGTCTGGCG 477
RESULT 2
LOCUS AX366054/c 1743 bp DNA linear PAT 15-FEB-2002
DEFINITION Sequence 447 from Patent WO0200174.
ACCESSION AX366054
VERSION AX366054.1 GI:18697498
KEYWORDS human.
ORGANISM Homo sapiens
SOURCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Wang, T., Wang, A., Skeiky, Y.A., Li, S.X., Kalos, M.D., Henderson, R.A.,
 McNeill, P.D., Fanger, N., Retter, M.W., Marnerakis, M., Fanger, G.R.,
 Vedvick, T.S., Carter, D., Watanabe, Y. and Peckham, D.W.
TITLE Compositions and methods for the therapy and diagnosis of lung
JOURNAL Patent: WO 0200174-A 447 03-JAN-2002;
FEATURES CORIXA CORPORATION (US)
 source Location/Qualifiers
 1. .1743
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
BASE COUNT 527 a 406 c 418 g 392 t
ORIGIN
 Query Match 100.0%; Score 22; DB 6; Length 1743;
 Best Local Similarity 100.0%; Pred. No. 5.6;
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 CTGCAAGGGGTTTGTCTGGCG 22
 Db 498 CTGCAAGGGGTTTGTCTGGCG 477
RESULT 3
LOCUS AX366057/c 1743 bp DNA linear PAT 15-FEB-2002
DEFINITION Sequence 450 from Patent WO0200174.
ACCESSION AX366057
VERSION AX366057.1 GI:18697500
KEYWORDS human.
ORGANISM Homo sapiens
SOURCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Wang, T., Wang, A., Skeiky, Y.A., Li, S.X., Kalos, M.D., Henderson, R.A.,
 McNeill, P.D., Fanger, N., Retter, M.W., Marnerakis, M., Fanger, G.R.,
 Vedvick, T.S., Carter, D., Watanabe, Y. and Peckham, D.W.
TITLE Compositions and methods for the therapy and diagnosis of lung
JOURNAL Patent: WO 0200174-A 450 03-JAN-2002;
FEATURES CORIXA CORPORATION (US)
 source Location/Qualifiers
 1. .1743
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
BASE COUNT 527 a 406 c 418 g 392 t
ORIGIN
 Query Match 100.0%; Score 22; DB 6; Length 1743;
 Best Local Similarity 100.0%; Pred. No. 5.6;
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 CTGCAAGGGGTTTGTCTGGCG 22
 Db 498 CTGCAAGGGGTTTGTCTGGCG 477
RESULT 4
LOCUS AX366035/c 1764 bp DNA linear PAT 15-FEB-2002
DEFINITION Sequence 428 from Patent WO0200174.
ACCESSION AX366035
VERSION AX366035.1 GI:18697484
KEYWORDS human.
ORGANISM Homo sapiens
SOURCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Wang, T., Wang, A., Skeiky, Y.A., Li, S.X., Kalos, M.D., Henderson, R.A.,
 McNeill, P.D., Fanger, N., Retter, M.W., Marnerakis, M., Fanger, G.R.,
 Vedvick, T.S., Carter, D., Watanabe, Y. and Peckham, D.W.
TITLE Compositions and methods for the therapy and diagnosis of lung
JOURNAL Patent: WO 0200174-A 428 03-JAN-2002;
FEATURES CORIXA CORPORATION (US)
 source Location/Qualifiers
 1. .1764
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
BASE COUNT 534 a 417 c 419 g 394 t
ORIGIN
 Query Match 100.0%; Score 22; DB 6; Length 1764;
 Best Local Similarity 100.0%; Pred. No. 5.6;
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 CTGCAAGGGGTTTGTCTGGCG 22
 Db 519 CTGCAAGGGGTTTGTCTGGCG 498
RESULT 5
LOCUS HSU76705/c 4155 bp mRNA linear PRI 26-JAN-1999
DEFINITION Human putative RNA binding protein Koc1 mRNA, complete cds.
ACCESSION U76705
VERSION U76705.1 GI:4098296
KEYWORDS Homo sapiens.
ORGANISM Homo sapiens
SOURCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 4155)
AUTHORS Mueller-Pillasch, F., Lacher, U. and Wallrapp, C.
TITLE Direct Submission
JOURNAL Submitted (30-OCT-1996) Innere Medizin I, University of Ulm,
 Robert-Koch-Str.8, Ulm 89081, Germany
FEATURES Location/Qualifiers
 source 1. .4155
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="7"


```

/map="7p11.5"
/cell_line="Patau9888s"
/tissue_type="pancreas"
251..1390
/functions="putative RNA binding protein"
/notes="Kh-domain containing protein; overexpressed in
cancer"
/codon_start=1
/product="Koc1"
/protein_id="A098297"
/db_xref="GI:4098297"
/translations="MNKLYGNLSNAAPSDLESIFDKAKIPVSGPFLVKTGYAFVDC
PDSWALKAEALSGKIELRGKIEVHSVPKQRIKQIRNIPPHLQWELDSLIV
QYGVSECEQVNDTSETAVVNTYSSKQARQALDKNGFLENFTLVKVAIPDENAA
QQNPLOQPRGRLGQSGSGSGSKQKPCDPLRLVPTQFVGAIGKEGATI
RNITKTQSKIDVHRKENAGAAEKSTILSTPGTSAACKSILEIMHKEAQDIKFTTE
IPLKILAHNNFVGRLLGKGRNLKIEQDITKTITISPLQELTLNPERITVKGVE
TCAKABEIMKKIRESYENDIASMNLQAHILFGLNMLGLFPFTSGMPPTSGPPSA
MTPPYQFQSESETVHQPALSGAIGKQGHQKLSRFAGASIKIAPAEAPDAK
VRNVIITGPPEAQKAGRIYKIKENFVSPKEVKEAHIRVPSFAAGRVIGKGGK
TVNELQNLSSAEVVPDQTPDENDQVVVKITGHFYACQVQAKRIQELITQVQKHOQQ
KALQSGPPQSRRK"
BASE COUNT 1277 a 830 c 851 g 1181 t 16 others
ORIGIN

Query Match 100.0%; Score 22; DB 9; Length 4155;
Best Local Similarity 100.0%; Pred. No. 5.3;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTGCAAGGGGTTTGTCTGGGCG 22
|||||
Db 748 CTGCAAGGGGTTTGTCTGGGCG 727

RESULT 6
AR171863/c
LOCUS AR171863 4159 bp DNA linear PAT 17-DEC-2001
DEFINITION Sequence 4 from patent US 6297364.
ACCESSION AR171863
VERSION AR171863.1 GI:17910813
KEYWORDS
SOURCE Unknown.
REFERENCE 1 (bases 1 to 4159)
AUTHORS Chen, Y.-T., Gure, A., Tsang, S., Stockert, E., Jager, E., Alexander, K.
and Old, L.J.
TITLE Isolated nucleic acid molecule encoding cancer associated antigen,
the antigen itself, and uses thereof
JOURNAL Patent: US 6297364-A 4 02-OCT-2001;
FEATURES Location/Qualifiers
SOURCE 1..4159
/organism="unknown"
BASE COUNT 1281 a 830 c 851 g 1181 t 16 others
ORIGIN

Query Match 100.0%; Score 22; DB 6; Length 4159;
Best Local Similarity 100.0%; Pred. No. 5.3;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTGCAAGGGGTTTGTCTGGGCG 22
|||||
Db 748 CTGCAAGGGGTTTGTCTGGGCG 727

RESULT 7
AF117108/c
LOCUS AF117108 4171 bp mRNA linear PRI 26-JAN-1999
DEFINITION Homo sapiens IGF-II mRNA-binding protein 3 (IMP-3) mRNA, complete
cds.
ACCESSION AF117108
VERSION AF117108.1 GI:4191611
KEYWORDS

```

```

SOURCE Homo sapiens.
ORGANISM Homo sapiens
Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 4171)
AUTHORS Nielsen, J., Christensen, J., Lykke-Andersen, J., Johnsen, A.H.,
Wewer, U.M., and Nielsen, F.C.
TITLE A family of insulin-like growth factor II mRNA-binding proteins
represses translation in late development
JOURNAL Mol. Cell. Biol. 19 (2), 1262-1270 (1999)
MEDLINE 99108099
PUBMED 9891060
REFERENCE 2 (bases 1 to 4171)
AUTHORS Nielsen, J., Christensen, J., Lykke-Andersen, J., Johnsen, A.H.,
Wewer, U.M., and Nielsen, F.C.
TITLE Direct Submission
JOURNAL Submitted (30-DEC-1998) Institute of Molecular Biology, University
of Copenhagen, Soelvgade 83H, Copenhagen DK-1307, Denmark
FEATURES Location/Qualifiers
SOURCE 1..4171
/organism="Homo sapiens"
/db_xref="taxon:9606"
1..4171
/gene="IMP-3"
267..2006
/gene="IMP-3"
/notes="binds multiple sites in IGF-II mRNA 5'UTR;
identical to KH-domain containing protein overexpressed in
cancer (koc)"
/codon_start=1
/product="IGF-II mRNA-binding protein 3"
/protein_id="A09828.1"
/db_xref="GI:4191612"
/translations="MNKLYGNLSNAAPSDLESIFDKAKIPVSGPFLVKTGYAFVDC
PDSWALKAEALSGKIELRGKIEVHSVPKQRIKQIRNIPPHLQWELDSLIV
QYGVSECEQVNDTSETAVVNTYSSKQARQALDKNGFLENFTLVKVAIPDENAA
QQNPLOQPRGRLGQSGSGSGSKQKPCDPLRLVPTQFVGAIGKEGATI
RNITKTQSKIDVHRKENAGAAEKSTILSTPGTSAACKSILEIMHKEAQDIKFTTE
IPLKILAHNNFVGRLLGKGRNLKIEQDITKTITISPLQELTLNPERITVKGVE
TCAKABEIMKKIRESYENDIASMNLQAHILFGLNMLGLFPFTSGMPPTSGPPSA
MTPPYQFQSESETVHQPALSGAIGKQGHQKLSRFAGASIKIAPAEAPDAK
VRNVIITGPPEAQKAGRIYKIKENFVSPKEVKEAHIRVPSFAAGRVIGKGGK
TVNELQNLSSAEVVPDQTPDENDQVVVKITGHFYACQVQAKRIQELITQVQKHOQQ
KALQSGPPQSRRK"
BASE COUNT 1292 a 822 c 853 g 1199 t 5 others
ORIGIN

Query Match 100.0%; Score 22; DB 9; Length 4171;
Best Local Similarity 100.0%; Pred. No. 5.3;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTGCAAGGGGTTTGTCTGGGCG 22
|||||
Db 764 CTGCAAGGGGTTTGTCTGGGCG 743

RESULT 8
AX333233/c
LOCUS AX333233 4181 bp DNA linear PAT 09-JAN-2002
DEFINITION Sequence 3742 from Patent WO0194629.
ACCESSION AX333233
VERSION AX333233.1 GI:18123867
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Young, P.E., Augustus, M., Carter, K.C., Ebner, R., Endress, G.,
Horrigan, S., Soppet, D.R., and Weaver, Z.
TITLE Cancer gene determination and therapeutic screening using signature
gene sets
JOURNAL Patent: WO 0194629-A 3742 13-DEC-2001;

```

2 (bases 1 to 4181)
Mueller-Pillasch, F., Lacher, U., Wallrapp, C., Micha, A.,
Zimmerhackl, F., Hameister, H., Varga, G., Fries, H., Buchler, M.,
Beger, H.G., Villa, M.R., Adler, G. and Gress, T.M.
Direct Submission
Submitted (11-APR-1997) Medizinische Klinik, Internal Medicine I,
Robert-Koch-Str.8, Ulm 89081, Germany
Location/Qualifiers
1. 4181
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="7"
/map="7p11.5"
/cell_line="Patu8988t"
/cell_type="pancreatic cancer"
1. 4181
/gene="koc"
251. 1990
/gene="koc"
/function="possible proliferation of cancer tissues"
/note="KH-domain containing protein overexpressed in
cancer"
/codon_start=1
/product="putative RNA binding protein KOC"
/protein_id="AAC35208.1"
/db_xref="GI:2105469"
/translation="MNKLYIGNLSENAAPSDLESIFPKDAKIPVSGPFLVKTGYAFVDC
PDSWALKAIEALSGKIELHGKPIEVEHSVPKQRIKQIRNIPHLOWELDSLLV
QVGVSCQVNTDSTAVVNTVYSSKQDQALDKINGFOLENFTLVAYIPDEWAA
CQNPLOQGRGLGORGSGROSGSGSKQPCDPLRLVLTQVFGAIIKGEKATI
RNITKQSKIDVHRKENAGAKESITLSTPEGTSAACKSILEINMHEKAQIDIKFTEE
IPUKIAHNNFVRLGKGRNLKIEQDITKITSPLQELTLNPNRTITVKGNSA
TCAAEEEMKIRESYENDIASMNLQALHLPGLNHLALGLPPTSGMPPPTGPPSA
MTPPYQEQSETETVHQFIPALSGAIIKQGHQIKLSRFAGASIKIAPAEAPAK
VRVIITGPPEAQFKAQRIYKIEENFVSKPEVKLEAHIRVPSFAGRVIGKGGK
TVNELQNLSSAEVVPRDQTPDENQVWVKITGHFYACQVAKRIQELITVQKHQQQ
KALQSGPPQSRKX"

BASE COUNT 1303 a 830 c 851 g 1181 t 16 others
ORIGIN

Query Match 100.0%; Score 22; DB 9; Length 4181;
Best Local Similarity 100.0%; Pred. No. 5.3;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTGCAAGGGGTTTCTGGGCG 22
|||||
Db 748 CTGCAAGGGGTTTCTGGGCG 727

RESULT 11
AX397963/c
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
Mueller, F., Gress, T. and Adler, G.
Medicament comprising a dna sequence, which codes for the
rna-binding koc protein, and comprising a koc protein or a dna
sequence of the koc promoter
Patent: WO 0220036-A 1 14-MAR-2002;
Mueller, Friederike (DE)
Location/Qualifiers
1. 4602
/organism="Homo sapiens"
/db_xref="taxon:9606"
672. 2411
/note="unnamed protein product"

REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
source
CDS

10:29:39 2003
Avalon Pharmaceuticals (US)
Location/Qualifiers
1. 4181
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 1303 a 830 c 851 g 1181 t 16 others
ORIGIN

Query Match 100.0%; Score 22; DB 6; Length 4181;
Best Local Similarity 100.0%; Pred. No. 5.3; 0; Indels 0; Gaps 0;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTGCAAGGGGTTTCTGGGCG 22
|||||
Db 748 CTGCAAGGGGTTTCTGGGCG 727

RESULT 9
AX365782/c
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
Wang, T., Wang, A., Skeiky, Y.A., Li, S.X., Kalos, M.D., Henderson, R.A.,
McNeill, P.D., Fanger, N., Retter, M.W., Marnerakis, M., Fanger, G.R.,
Vedwick, T.S., Carter, D., Watanabe, Y. and Peckham, D.W.
Compositions and methods for the therapy and diagnosis of lung
cancer
Patent: WO 0200174-A 175 03-JAN-2002;
CORIXA CORPORATION (US)
Location/Qualifiers
1. 4181
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 1303 a 830 c 851 g 1181 t 16 others
ORIGIN

Query Match 100.0%; Score 22; DB 6; Length 4181;
Best Local Similarity 100.0%; Pred. No. 5.3; 0; Indels 0; Gaps 0;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTGCAAGGGGTTTCTGGGCG 22
|||||
Db 748 CTGCAAGGGGTTTCTGGGCG 727

RESULT 10
HSU97188/c
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 4181)
Mueller-Pillasch, F., Lacher, U., Wallrapp, C., Micha, A.,
Zimmerhackl, F., Hameister, H., Varga, G., Fries, H., Buchler, M.,
Beger, H.G., Villa, M.R., Adler, G. and Gress, T.M.
Cloning of a gene highly overexpressed in cancer coding for a novel
KH-domain containing protein
OncoGene 14 (22), 2729-2733 (1997)
97322117
9178771

```

/codon_start=1
/protein_id="CAD32603.1"
/db_xref="GI:21360819"
/translation="MKKLYICNLSNAAPSDLESIFKDAKIPVSGPFLVKTGVAFVDC
PDSWALKALEALUSGKIELHGHKPIEVHVSVPKQRIKQLQIRNIPHLQWEVLDSLLV
QYGVSECEQVNTDSTAVVNVVTSKQQAQALDKLNGFLENFTLKVAYIPDEMAA
QNPLOQRRGLGSGRSRQSGSVSKQKCDPLRLVPTQVGAIGKEGATI
RNTTKQTQSDIVHRNAGAAEKSTILSTPGTSAACKSILEIMHKEAQDIKTEE
IPUKILAHNNFVGLIKEGRNLIKTEQDQDKITISPLQELTLNPERITIVKNGVE
TCABEEIMKKIRSEYENDIASMNLOAHILFGLNUNALGLFPPTSGPPPTSGPPSA
MTPTPQFQSESETVHQFIPALSVGAILIGKQGHKIKQLSRFAGASIKTAPAEAPDAK
TVNELITGPEAFQKAGRIYKIKENFVSPKBEKLEAHIRVSPFAGRIVIGKGGK
VNWLNLSAEAVVPRDQTPDENDQVVKITGHFYACQVAKIQEILTVQKQHQQ
KALQSGPPQSRRT"
BASE COUNT 1463 a 907 c 918 g 1298 t 16 others
ORIGIN
Query Match 100.0%; Score 22; DB 6; Length 4602;
Best Local Similarity 100.0%; Pred. No. 5.2;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 CTGCAAGGGGTTTGTCTGGGCG 22
|||||
Db 1169 CTGCAAGGGGTTTGTCTGGGCG 1148

RESULT 12
AC021876 104668 bp DNA linear PRI 04-JUN-2002
LOCUS Homo sapiens BAC clone GS1-117B4 from 7, complete sequence.
DEFINITION AC021876
ACCESSION AC021876
VERSION AC021876.5 GI:13446341
KEYWORDS HTG.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 104668)
Sulston,J.E. and Waterston,R.
Toward a complete human genome sequence
Genome Res. 8 (11), 1097-1108 (1998)
99063792
9847074
PUBMED
2 (bases 1 to 104668)
Du,F., Stoneking,T. and Moeller,D.
The sequence of Homo sapiens BAC clone GS1-117B4
Unpublished (2001)
3 (bases 1 to 104668)
Waterston,R.H.
Direct Submission
Submitted (21-JAN-2000) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
4 (bases 1 to 104668)
Waterston,R.H.
Direct Submission
Submitted (03-AUG-2000) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
5 (bases 1 to 104668)
Waterston,R.H.
Direct Submission
Submitted (10-SEP-2000) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
6 (bases 1 to 104668)
Waterston,R.
Direct Submission
Submitted (30-SEP-2000) Department of Genetics, Washington
University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
7 (bases 1 to 104668)
Waterston,R.H.
Direct Submission

```

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Submitted (25-MAR-2001) Genome Sequencing Center, Washington University School of Medicine, 4444 Forest Park Parkway, St. Louis, MO 63108, USA
 8 (bases 1 to 104668)
 Waterston,R.
 Direct Submission
 Submitted (01-MAR-2002) Department of Genetics, Washington University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
 9 (bases 1 to 104668)
 Waterston,R.
 Direct Submission
 Submitted (04-JUN-2002) Department of Genetics, Washington University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
 On Mar 25, 2001 this sequence version replaced gi:9665195.
 ----- Genome Center
 Center: Washington University Genome Sequencing Center
 Center code: WUGSC
 Web site: <http://genome.wustl.edu/gsc>
 Contact: sapiens@watson.wustl.edu
 ----- Summary Statistics

 Center project name: H_GS117B04

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted: all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:

The sequence of this clone was established as part of a mapping and sequencing collaboration between the NHGRI Chromosome 7 Mapping Project (Eric D. Green, Director), John D. McPherson in the Department of Genetics (Washington University), and the Washington University Genome Sequencing Center. For additional information about the map position of this sequence, see <http://www.nhgri.nih.gov/DIR/GRB/CHR7>, send mailto:egreen@nhgri.nih.gov, or see <http://genome.wustl.edu/gsc>

SOURCE INFORMATION:

This clone is from the first BAC library from Genome Systems, Inc. (<http://www.genomesystems.com>).
 Cell line: lymphoblastoid
 Haplotypes: two
 VECTOR: pBelOBAC
 Selection: chloramphenicol

NEIGHBORING SEQUENCE INFORMATION:

The clone sequenced to the left is CTA-271G13, 200 bp overlap; the clone sequenced to the right is AC023375. Actual start of this clone is at base position 102112 of CTA-271G13; actual end is at base position 104668 of GS1-117B4.

The sequence from 80185 to 80225 was derived from PCR product of GS1-117B4 BAC DNA.

FEATURES

source

1. .104668
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="7"
 /map="7"
 /clone="GS1-117B4"
 /clone_lib="GSBAC1"
 1042..1339
 /rpt_family="Alu"
 1733..2070

repeat_region

misc_feature

12928. .13050
/note="similar to EST BF241103 (NID:g11155028)"
12945. .13050
/note="similar to EST AV612377 (NID:g9748047)"
13115. .13172
/rpt_family="L2"
13571. .13877
/rpt_family="Alu"
14561. .14722

Query Match 100.0%; Score 22; DB 9; Length 104668;
Best Local Similarity 100.0%; Pred. No. 4.1;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0

Qy 1 CTCGACGGGTTTGTCTGGGCG 22
|||||
Db 18478 CTCGACGGGTTTGTCTGGGCG 18499
|||||

RESULT 13
AC023375 LOCUS
DEFINITION Homo sapiens chromosome 7 clone RP11-571M6 map 7, WORKING DRAFT
SEQUENCE, 37 unordered pieces.
AC023375 AC023375.2 GI:7209933
VERSION HTG; HTGS_PHASE1; HTGS_DRAFT.
KEYWORDS human.
SOURCE Homo sapiens
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 158105)
AUTHORS Birren,B., Linton,L., Nusbaum,C. and Lander,E.
TITLE Homo sapiens chromosome 7, clone RP11-571M6
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 158105)
AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham H., Allen,N.,
Anderson,S., Baldwin,J., Barna,N., Beda,F., Boguslavsky,L.,
Boukhalter,B., Brown,A., Burkett,G., Campopiano,A., Castle,A.,
Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cooke,P.,
DeArellano,K., Dewar,K., Dodge,S., Domino,M., Doyle,M.,
Fenster,J., Ferreira,P., FitzHugh,W., Forrest,C., Gage,D.,
Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L.,
Grand-Pierre,N., Grant,G., Hagos,B., Heaford,A., Horton,L.,
Howard,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A.,
Klein,J., Landers,T., Larcocque,K., Lehoczeky,J., Levine,R.,
Lieu,C., Liu,G., Locke,K., Macdonald,P., Marquis,N., McCarthy,M.,
McEwan,P., McGurk,A., Miranda,C., Mieng,V., Morrow,J., Naylor,J.,
Neuman,L., Mihova,T., Miranda,C., Mieng,V., Morrow,J., Naylor,J.,
Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D., Olivari,T.M.,
Peterson,K., Pierre,N., Pisan,C., Pollara,V., Raymond,C.,
Riley,R., Rogov,P., Rothman,D., Roy,A., Santos,R., Schauer,S.,
Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J., Tirrell,A.,
Travers,M., Trigglio,J., Vassiliev,H., Viel,R., Vo,A., Wilson,B.,
Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J., Zimmer,A. and
Zody,M.

TITLE Direct Submission
JOURNAL Submitted (14-FEB-2000) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
REFERENCE 3 (bases 1 to 158105)
AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham H., Allen,N.,
Anderson,S., Baldwin,J., Barna,N., Bastien,V., Beda,F.,
Boguslavsky,L., Boukhalter,B., Brown,A., Burkett,G.,
Campopiano,A., Castle,A., Choepel,Y., Colangelo,M., Collins,S.,
Collymore,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S.,
Dodge,S., Domino,M., Doyle,M., Ferreira,P., FitzHugh,W., Gage,D.,
Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L.,
Grand-Pierre,N., Grant,G., Hagos,B., Heaford,A., Horton,L.,
Howard,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A.,
Klein,J., Larcocque,K., Lamazares,R., Landers,T., Lehoczeky,J.,
Levine,R., Lieu,C., Liu,G., Locke,K., Macdonald,P., Marquis,N.,
McCarthy,M., McEwan,P., McGurk,A., McKernan,K., McPheeters,R.,

Meldrim, J., Meneus, L., Mihova, T., Miranda, C., Mlenga, V., Morrow, J., Murphy, T., Naylor, J., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neil, D., Oliver, T.M., Oliver, J., Peterson, K., Pierre, N., Piani, C., Pollara, V., Raymond, C., Riley, R., Rogov, P., Rothman, D., Roy, A., Santos, R., Schauer, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J., Tirrell, A., Travers, M., Trigilio, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J., Young, G., Zainoun, J., Zimmer, A. and Zody, M.

Direct Submission
Submitted (24-AUG-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
On Mar 8, 2000 this sequence version replaced gi:6970513.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

TITLE JOURNAL COMMENT

Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR

Web site: <http://www.seq.wi.mit.edu>

Contact: sequence_submissions@genome.wi.mit.edu

Project Information

Center project name: L4227

Center clone name: 571.M.6

Summary Statistics

Sequencing vector: M13; M7815; 100% of reads

Chemistry: Dye-terminator Big Dye; 100% of reads

Assembly program: Phrap; version 0.960731

Consensus quality: 133946 bases at least Q40

Consensus quality: 143828 bases at least Q30

Consensus quality: 149412 bases at least Q20

Insert size: 174000; agarose-1p

Quality coverage: 3.2 in Q20 bases; agarose-1p

Quality coverage: 3.6 in Q20 bases; sum-of-contigs

NOTE: This is a 'working draft' sequence. It currently consists of 37 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

1 1036: contig of 1036 bp in length
1037 1136: gap of 100 bp
1137 2213: contig of 1077 bp in length
2214 2313: gap of 100 bp
2314 3566: contig of 1253 bp in length
3567 3666: gap of 100 bp
3667 4804: contig of 1138 bp in length
4805 4904: gap of 100 bp
4905 6404: contig of 1500 bp in length
6405 6504: gap of 100 bp
6505 8366: contig of 1862 bp in length
8367 8466: gap of 100 bp
8467 10126: contig of 1660 bp in length
10127 10226: gap of 100 bp
10227 11460: contig of 1234 bp in length
11461 11560: gap of 100 bp
11561 13697: contig of 2137 bp in length
13698 13797: gap of 100 bp
13798 15334: contig of 1537 bp in length
15335 15434: gap of 100 bp
15435 16434: contig of 1000 bp in length
16435 16534: gap of 100 bp
16535 18900: contig of 2366 bp in length
18901 19000: gap of 100 bp
19001 20918: contig of 1918 bp in length
20919 21018: gap of 100 bp
21019 23110: contig of 2092 bp in length
23111 23210: gap of 100 bp
23211 25903: contig of 2693 bp in length

25904 26003: gap of 100 bp
26004 29042: contig of 3039 bp in length
29043 29142: gap of 100 bp
29143 32196: contig of 3054 bp in length
32197 32296: gap of 100 bp
32297 35269: contig of 2973 bp in length
35270 35369: gap of 100 bp
35370 39031: contig of 3662 bp in length
39032 39131: gap of 100 bp
39132 42030: contig of 2899 bp in length
42031 42130: gap of 100 bp
42131 46601: contig of 4471 bp in length
46602 46701: gap of 100 bp
46702 51387: contig of 4686 bp in length
51388 51487: gap of 100 bp
51488 55227: contig of 3740 bp in length
55228 55327: gap of 100 bp
55328 59146: contig of 3819 bp in length
59147 59246: gap of 100 bp
59247 62533: contig of 3293 bp in length
62540 62633: gap of 100 bp
62640 66443: contig of 3804 bp in length
66444 66543: gap of 100 bp
66544 72068: contig of 5525 bp in length
72069 72168: gap of 100 bp
72169 78427: contig of 6259 bp in length
78428 78527: gap of 100 bp
78528 83366: contig of 4839 bp in length
83367 83466: gap of 100 bp
83467 89726: contig of 6260 bp in length
89727 89826: gap of 100 bp
89827 97979: contig of 8153 bp in length
97980 98079: gap of 100 bp
98080 106168: contig of 8089 bp in length
106169 106268: gap of 100 bp
106269 114387: contig of 8119 bp in length
114388 114487: gap of 100 bp
114488 121719: contig of 7232 bp in length
121720 121819: gap of 100 bp
121820 130088: contig of 8269 bp in length
130089 130188: gap of 100 bp
130189 139207: contig of 9019 bp in length
139208 139307: gap of 100 bp
139308 158105: contig of 18798 bp in length.

FEATURES

Location/Qualifiers

1..158105
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="7"
/map="7"
/clone="RP11-571M6"
/clone_lib="RPC1-11 Human Male BAC"

misc_feature
1..1036
/note="assembly_fragment"
misc_feature
1137..2213
/note="assembly_fragment"
misc_feature
2314..3566
/note="assembly_fragment"
misc_feature
3667..4804
/note="assembly_fragment"
misc_feature
4905..6404
/note="assembly_fragment"
misc_feature
6505..8366
/note="assembly_fragment"
misc_feature
8467..10126
/note="assembly_fragment"
misc_feature
10227..11460
/note="assembly_fragment"
misc_feature
11561..13697
/note="assembly_fragment"
misc_feature
13798..15334
/note="assembly_fragment"
misc_feature
15435..16434

```

/misc_feature
/note="assembly_fragment"
16535.18900

Query Match
Best Local Similarity 100.0%; Score 22; DB 2; Length 158105;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTGCAAGGGGTTTCTGGGCG 22
|||||
Db 148579 CTGCAAGGGGTTTCTGGGCG 148600

RESULT 14
AC090069 98951 bp DNA linear HTG 20-APR-2001
LOCUS Homo sapiens chromosome 5 clone CTD-2162G15, WORKING DRAFT
DEFINITION SEQUENCE, 7 unordered pieces.
AC090069
AC090069.1 GI:12830173
HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ACTIVEFIN.
VERSION
KEYWORDS
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 98951)
AUTHORS DOE Joint Genome Institute.
TITLE Sequencing of Human Chromosome 5
JOURNAL Unpublished
REFERENCE
2 (bases 1 to 98951)
AUTHORS DOE Joint Genome Institute.
TITLE Direct Submission
JOURNAL Submitted (14-FEB-2001) Production Sequencing Facility, DOE Joint
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA

-----Genome Center
Center: Joint Genome Institute
Center Code: JGI
Web site: http://www.jgi.doe.gov

-----
Project Information
Center Project Name: 689916
Center clone name: CITB-H1_2162G15

-----
Summary Statistics
Consensus quality: 91863 bases at least Q40
Consensus quality: 96729 bases at least Q30
Consensus quality: 97483 bases at least Q20
Estimated insert size: 103000; agarose-fp estimation
Estimated insert size: 98351; sum-of-contigs estimation
Quality coverage: 5.97 in Q20 bases; agarose-fp estimation
Quality coverage: 5.97 in Q20 bases; sum-of-contigs estimation.
* NOTE: This is a 'working draft' sequence. It currently
* consists of 7 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 1640: contig of 1640 bp in length
* 1641 1740: gap of unknown length
* 1741 4817: contig of 3077 bp in length
* 4818 4917: gap of unknown length
* 4918 10479: contig of 5562 bp in length
* 10480 10579: gap of unknown length
* 10580 17185: contig of 6606 bp in length
* 17186 17285: gap of unknown length
* 17286 27136: contig of 9851 bp in length
* 27137 27236: gap of unknown length
* 27237 53346: contig of 26110 bp in length
* 53347 53447: gap of unknown length
* 53447 98951: contig of 45505 bp in length.

FEATURES
Location/Qualifiers
1..98951

```

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="5"
/clone="CTD-2162G15"
/clone_lib="Caltech human BAC library D"
28584 a 22060 c 20767 g 26935 t 605 others

BASE COUNT
ORIGIN

Query Match
Best Local Similarity 85.5%; Score 18.8; DB 2; Length 98951;
Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CTGCAAGGGGTTTCTGGGCG 22
|||||
Db 10782 CTGCAAGGGGTTTCTGTGAG 10803

RESULT 15
AC104165 200956 bp DNA linear PRI 27-JUN-2002
LOCUS Homo sapiens chromosome 3 clone RP11-68104, complete sequence.
DEFINITION AC104165 AC068625
AC068625
AC104165.2 GI:21617715
VERSION
KEYWORDS HTG.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 200956)
AUTHORS Kaul,R.K., Olson,M.V., Zhou,Y., James,R.A., Rouse,G., Wu,Z.,
Saenphimmachak,C., Phelps,K.A., Buckley,D., Kibukawa,M., Raymond,C.
and Haugen,E.D.
TITLE Direct Submission
JOURNAL Unpublished
REFERENCE
2 (bases 1 to 200956)
AUTHORS Kaul,R.K., Olson,M.V., Raymond,C. and Haugen,E.D.
TITLE Direct Submission
JOURNAL Submitted (05-DEC-2001) Genome Center, University of Washington,
Box 352145, Seattle, WA 98195, USA
REFERENCE
3 (bases 1 to 200956)
AUTHORS Kaul,R.K., Olson,M.V., Zhou,Y., James,R.A., Rouse,G., Wu,Z.,
Saenphimmachak,C., Phelps,K.A., Buckley,D., Kibukawa,M., Raymond,C.
and Haugen,E.D.
TITLE Direct Submission
JOURNAL Submitted (27-JUN-2002) Genome Center, University of Washington,
Box 352145, Seattle, WA 98195, USA
On Jun 27, 2002 this sequence version replaced gi:17352429.
-----Genome Center
Center: University of Washington Genome Center
Center Code: UWGC
Web site: http://www.genome.washington.edu
Contact: uwgctgs@u.washington.edu
Drafting Center: WUGSC
-----Project Information
Center project name: chr-3
Center clone name: RP11-68104 (bc0528)
-----Summary Statistics
Sequencing vector: unknown; 42% of reads
Sequencing vector: plasmid; 58% of reads
Chemistry: Dye-terminator ET; 80% of reads
Chemistry: Dye-terminator Big Dye; 20% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 200732 bases at least Q40
Consensus quality: 200909 bases at least Q30
Consensus quality: 200947 bases at least Q20
Insert size: 200956; sum-of-contigs
Quality coverage: 10.4x in Q20 bases; sum-of-contigs
-----
Overlapping Sequences:
5': RP11-348P10 (UWGC:bc0377) AC124045
3': RP11-578F5 (UWGC:bc0477) AC105902, 95352-bp overlap
-----
Sequence Quality Assessment:

```

This entry has been annotated with sequence quality estimates computed by the Phrap assembly program. All manually edited bases have been reduced to quality zero. Quality levels above 40 are expected to have less than 1 error in 10,000 bp.

Base-by-base quality values are not generally visible from the GenBank flat file format but are available as part of this entry's ASN.1 file.

This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., Phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest.

Sequence Validation:

This sequence has been validated by Multiple Complete Digest fingerprinting. Comparison of the experimentally derived digest fragments with sequence-predicted fragments is given below. The electronically-digested sequence consists of both insert and vector, in order to accurately represent the entire circular BAC. Small fragments below a variable cutoff (approximately 400-800 bp) are not resolved in the fingerprint and hence do not appear in the table. There are no significant remaining discrepancies between the experimental and predicted values. Uniquely ordered fragments are separated by dashed lines.

BgIII			EcoRI			HindIII		
SeqDerMap	FngPrnt	SeqDerMap	FngPrnt	SeqDerMap	FngPrnt	SeqDerMap	FngPrnt	SeqDerMap
5328	5288	8696	8571	3534	3560			
2067	2057	6	<800	6382	6318			
6846	6934	6068	6154	512	<800			
144	<800	11400	11207	449	<800			
1459	1439	96	<800	5736	5646			
70	<800	4100	4106	2978	3031			
3690	3725	779	760	8146	8133			
102	<800	1221	1170	4465	4477			
511	<800	278	<800	2239	2257			
1903	1919	8458	8571	45	<800			
5124	5028	12014	11879	4368	4477			
367	<800	645	<800	286	<800			
7100	6934	1177	1170	4633	4784			
5890	5889	153	<800	1462	1405			
1814	1816	1993	1932	2058	2110			
125	<800	3153	3186	745	<800			
43	<800	132	<800	1321	1243			
117	<800	4939	4965	3528	3560			
1955	1919	466	<800	1640	1662			
3477	3529	1311	1279	8779	8765			

4899	5028	3183	3186	5089	4991
774	799	16952	16678	2571	2631
10749	10616	1629	1629	2440	2446
1620	1602	7261	7427	1682	1662
8868	8810	7837	7939	1660	1662
5884	5889	3940	4106	4912	4991
1521	1529	756	760	2470	2446
771	799	2395	2496	8892	8765
11706	11675	5310	5399	1257	1243
369	<800	21139	21498	10114	9869
5177	5288	4082	4106	14019	14339
2916	2950	7866	7939	3580	3560
137	<800	7411	7427	3022	3031
2406	2347	3378	3366	8296	8133
3857	3725	7272	7427	3319	3311
5249	5288	3059	3186	1694	1662
5343	5288	8214	8571	1747	1662
6819	6934	5836	5849	2370	2446
2335	2347	5394	5399	4817	4784
569	<800	419	<800	1139	1073
135	<800	3260	3186	3946	3909
3825	3725	2082	2097	12931	12902
6741	6934	8568	8571	2109	2110
1472	1439	1681	1629	364	<800
3707	3725	741	760	78	<800
703	<800	2905	3005	254	<800
207	<800			6129	6318
5047	5028			1674	1662
1134	1133			1250	1243
1070	1069			4768	4784
4665	4675			1083	1073
1121	1133			1023	1073
6920	6934			2602	2631
3317	3264			2449	2446
1683	1697			1249	1243
3525	3529			7228	7182

us-09-270-437d-10.rge

Mon Apr 21 10:29:39 2003

Query Match 85.5%; Score 18.8; DB 9; Length 200956;
 Best Local Similarity 90.9%; Pred. No. 1.5e+02;
 Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CTGCAAGGGGTTTGTCTGGCG 22
 |||||
 Db 27941 CTGCCAGGGGTTTGTCTGGTG 27962

Search completed: April 18, 2003, 06:17:22
 Job time : 385 secs

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 18, 2003, 04:12:35 ; Search time 98 Seconds
(without alignments)
505.551 Million cell updates/sec

Title: US-09-270-437D-10

Perfect score: 22
Sequence: 1 ctgcaagggttttctgtggcg 22

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 112599159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_101002.*
1: /SID22/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*
2: /SID22/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*
3: /SID22/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*
4: /SID22/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*
5: /SID22/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*
6: /SID22/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*
7: /SID22/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*
8: /SID22/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*
9: /SID22/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*
10: /SID22/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*
11: /SID22/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*
12: /SID22/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*
13: /SID22/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*
14: /SID22/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*
15: /SID22/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*
16: /SID22/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*
17: /SID22/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*
18: /SID22/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*
19: /SID22/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*
20: /SID22/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*
21: /SID22/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*
22: /SID22/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*
23: /SID22/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*
24: /SID22/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	22	100.0	22	21	AAZ36156
2	22	100.0	282	22	AAZ36156
3	22	100.0	282	22	AAZ36156
4	22	100.0	588	22	AAZ36156
5	22	100.0	588	22	AAZ36156
6	22	100.0	1740	21	AAZ36156
7	22	100.0	1740	21	AAZ36156
8	22	100.0	1743	24	ABL49299
9	22	100.0	1743	24	ABL49299

C 10	22	100.0	1764	24	ABL49283	Human lung tumour
C 11	22	100.0	4159	21	AAZ36150	DNA encoding cancer
C 12	22	100.0	4181	21	AAZ36150	Human lung cancer-
C 13	22	100.0	4181	21	AAZ36150	Lung cancer relate
C 14	22	100.0	4181	21	AAZ36150	Human lung tumour
C 15	22	100.0	4601	24	ABA99558	Human KOC DNA. Ho
C 16	17.4	73.1	861	22	AAZ36156	Human neuroblastom
C 17	17.2	78.2	255	22	AAZ36156	DNA encoding novel
C 18	17.2	78.2	1206	21	AAZ36156	Drosophila melanog
C 19	17.2	78.2	1206	21	AAZ36156	Bacterium 2412.1 f
C 20	17.2	78.2	1985	23	AAZ36156	Oesophagus cancer relat
C 21	17.2	78.2	4264	23	AAZ36156	Human genomic DNA
C 22	17.2	78.2	18776	23	AAZ36156	Corn tassal-derive
C 23	17.2	78.2	24494	21	AAZ36156	Human colon cancer
C 24	17.2	78.2	302250	24	ABL67703	Human prostate exp
C 25	16.4	74.5	568	24	ABN64702	Human prostate exp
C 26	16.2	73.6	296	24	ABL72849	Human immune/haema
C 27	16.2	73.6	300	21	AAA00921	DNA encoding novel
C 28	16.2	73.6	462	23	ABV55571	Human gene express
C 29	16.2	73.6	514	23	ABV54482	Human colon cancer
C 30	16.2	73.6	533	22	AAK77132	Corn acid triacylg
C 31	16.2	73.6	561	23	AAZ36156	Streptococcus poly
C 32	16.2	73.6	782	20	AAZ36156	Drosophila melanog
C 33	16.2	73.6	784	21	AAA02465	Drosophila melanog
C 34	16.2	73.6	859	21	AAZ36156	E. coli strain DSM
C 35	16.2	73.6	1563	24	ABN67088	Plasmid pMD2 DNA.
C 36	16.2	73.6	2286	23	ABU02757	Human Kruppel-like
C 37	16.2	73.6	4739	23	ABU02757	Human Kruppel-like
C 38	16.2	73.6	4750	23	ABU02757	Human nervous syst
C 39	16.2	73.6	5547	20	AAZ36156	Human acetyl-Coenz
C 40	16.2	73.6	5552	19	AAZ36156	
C 41	16.2	73.6	6232	24	AAZ36156	
C 42	16.2	73.6	6232	24	AAZ36156	
C 43	16.2	73.6	9722	22	ABA15166	
C 44	16.2	73.6	14855	22	ABA15166	
C 45	16.2	73.6	22028	24	ABA93286	

ALIGNMENTS

RESULT 1
ID AAZ36156 standard; DNA; 22 BP.
XX AAZ36156;
AC AAZ36156;
XX
DT 11-FEB-2000 (first entry)
XX
DE PCR primer for DNA encoding cancer associated antigen KOC-1.
XX
KW Cancer associated antigen; KOC-1; cancer; vaccine; CT7; PCR primer; ss.
OS Synthetic.
OS Homo sapiens.
PN WO9954738-A1.
XX
PD 28-OCT-1999.
XX
PF 16-MAR-1999; 99WO-US05766.
XX
PR 17-APR-1998; 98US-0061709.
XX
XX (LUDW-) LUDWIG INST CANCER RES.
PA Chen Y, Gure A, Tsang S, Stockert E, Jager E, Knuth A, Old LJ;
PI WPI; 2000-013284/01.
DR Nucleotides representing cancer-associated genes, used to develop
PT products for the diagnosis, monitoring and treatment of cancers -

```

SQ      Sequence 282 BP; 77 A; 79 C; 70 G; 56 T; 0 other;

Query Match      100.0%; Score 22; DB 22; Length 282;
Best Local Similarity 100.0%; Pred. No. 0.72;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

AC DAT56378:

DT	17-OCT-2001 (first entry)
XX	Probe #25064 used to measure gene expression in human placenta sample.
DE	
XX	Probe; microarray; human; placenta; antenatal diagnosis;
KW	genetic disorder; ss.

NOZ00137272 111.
XX
PD 09-AUG-2001.

FD-302 (Rev. 1-25-60)

04-FEB-2000: 2000US-0180312.

30 447 3000: 3000TC-0608408

PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.

[illegible]XX
WP: 2001-488897/53.

Human genome-derived single exon nucleic acid probes useful for measuring gene expression in human placenta - including gene expression in human placenta -

XX claim 25: SEQ ID No 25064: 654pp; English.
XX analyzing gene expression in human placenta.

XX the present invention relates to single exon nucleic acid probes (SENP).

CC The present invention relates to single exon nucleic acid probes (SENPs).
 CC The present sequence is one such probe. The probes are useful for
 CC

CC The present sequence is one such probe. The probes are useful for
CC producing a microarray for predicting, measuring and displaying gene

CC producing a microarray for predicting, measuring and displaying some
CC expression in samples derived from human placenta. The probes are useful
CC in the diagnosis of human genetic disorders.

CC expression in samples derived from human placenta, and for
CC for antenatal diagnosis of human genetic disorders.

CC LOI antenatal diagnosis of human g
XX
SO Sequence 282 BP: 77 A; 79 C; 70 G; 56 T; 0 other;

Query	Match	Score 22;	DB 22;	Length 282;
SQ Sequence 282 BF; 77 A; 75 C; 70 G; 35 T; 6 U; 100.0%				

Query Match	100.0%;	Score
Best Local Similarity	100.0%;	Pred. No. 0.72;
Matches 22: Conservative	0;	Mismatches 0; Indels

```
Matches 22; Conservative 0; Mismatches 0; Inliers 0; Cops 0
```

76

RESULT 4
AAK37520/c
ID AAK37520 standard; DNA; 588 BP.

```

XX AC AAK37520;
XX AC
XX DT 06-NOV-2001 (first entry)
XX PR
XX DE Human bone marrow expressed single exon probe SEQ ID NO: 12077.
XX KW Human; bone marrow expressed exon; gene expression analysis; probe;
XX KW microarray; cancer; leukaemia; lymphoma; myeloma; ss.
XX OS Homo sapiens.
XX PN WO200157276-A2.
XX PD
XX PF 09-AUG-2001.
XX PR
XX PR 30-JAN-2001; 2001WO-US00668.
XX PR 04-FEB-2000; 2000US-0180312.
XX PR 26-MAY-2000; 2000US-0207456.
XX PR 30-JUN-2000; 2000US-0608408.
XX PR 03-AUG-2000; 2000US-0632366.
XX PR 21-SEP-2000; 2000US-0234687.
XX PR 27-SEP-2000; 2000US-0236359.
XX PR 04-OCT-2000; 2000GB-0024263.
XX PA (MOLE-) MOLECULAR DYNAMICS INC.
XX PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX PD
XX PF WPI; 2001-488900/53.
XX PR Human genome-derived single exon nucleic acid probes useful for
XX PT analyzing gene expression in human bone marrow -
XX PS Example 4; SEQ ID NO: 12077; 659pp + Sequence Listing; English.
XX CC The present invention provides a number of single exon nucleic acid
XX CC probes which are derived from genomic sequences expressed in the human
XX CC bone marrow. They can be used to measure gene expression in bone marrow
XX CC samples, which may enable the improved diagnosis and treatment of cancers
XX CC such as lymphoma, leukaemia and myeloma. The present sequence is one of
XX CC the probes of the invention.
XX SQ Sequence 588 BP; 156 A; 139 C; 143 G; 150 T; 0 other;

Query Match 100.0%; Score 22; DB 22; Length 588;
Best Local Similarity 100.0%; Pred. No. 0.78;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTGCAAGGGGTTTGTGGGCG 22
Db 117 CTGCAAGGGGTTTGTGGGCG 96

RESULT 5
AAI43398/c
ID AAI43398 standard; DNA; 588 BP.
XX AC
XX AC AAI43398;
XX DT
XX DT 17-OCT-2001 (first entry)
XX DE
XX DE Probe #12084 used to measure gene expression in human placenta sample.
XX KW Probe; microarray; human; placenta; antenatal diagnosis;
XX KW genetic disorder; ss.
XX OS Homo sapiens.
XX PN WO200157272-A2.
XX PD
XX PF 09-AUG-2001.

```

```

XX PF 30-JAN-2001; 2001WO-US00663.
XX PR
XX PR 04-FEB-2000; 2000US-0180312.
XX PR 26-MAY-2000; 2000US-0207456.
XX PR 30-JUN-2000; 2000US-0608408.
XX PR 03-AUG-2000; 2000US-0632366.
XX PR 21-SEP-2000; 2000US-0234687.
XX PR 27-SEP-2000; 2000US-0236359.
XX PR 04-OCT-2000; 2000GB-0024263.
XX PA (MOLE-) MOLECULAR DYNAMICS INC.
XX PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX PD
XX PF WPI; 2001-488997/53.
XX PR Human genome-derived single exon nucleic acid probes useful for
XX PT analyzing gene expression in human placenta -
XX PS Claim 25; SEQ ID No 12084; 654pp; English.
XX CC The present invention relates to single exon nucleic acid probes (SENP).
XX CC The present sequence is one such probe. The probes are useful for
XX CC producing a microarray for predicting, measuring and displaying gene
XX CC expression in samples derived from human placenta. The probes are useful
XX CC for antenatal diagnosis of human genetic disorders.
XX SQ Sequence 588 BP; 156 A; 139 C; 143 G; 150 T; 0 other;

Query Match 100.0%; Score 22; DB 22; Length 588;
Best Local Similarity 100.0%; Pred. No. 0.78;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTGCAAGGGGTTTGTGGGCG 22
Db 117 CTGCAAGGGGTTTGTGGGCG 96

RESULT 6
AAC66035/c
ID AAC66035 standard; cDNA; 1740 BP.
XX AC
XX AC AAC66035;
XX DT
XX DT 21-FEB-2001 (first entry)
XX DE
XX DE Human lung cancer-associated cDNA antigen L523S.
XX KW Lung cancer; therapy; treatment; human; tumor; immunogenic; cytostatic;
XX KW vaccine; detection; ss.
XX OS Homo sapiens.
XX PN WO200061612-A2.
XX PD
XX PD 19-OCT-2000.
XX PF
XX PF 03-APR-2000; 2000WO-US08896.
XX PR
XX PR 02-APR-1999; 99US-0285479.
XX PR 17-DEC-1999; 99US-0466396.
XX PR 30-DEC-1999; 99US-0476496.
XX PR 10-JAN-2000; 2000US-0480884.
XX PR 22-FEB-2000; 2000US-0510376.
XX PA (CORI-) CORIXA CORP.
XX PI Wang T, Fan L;
XX PD
XX PD WPI; 2000-628399/60.
XX DR P-PSDB; AAB11365.
XX PD

```


SQL Sequence 1743 BP; 527 A; 406 C; 418 G; 392 T; 0 other;

Query Match 100.0%; Score 22; DB 24; Length 1743;
Best Local Similarity 100.0%; Pred. No. 0.9;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTGCAAGGGGTTTGTCTGGCG 22
| | | | | | | | | | | | | | | | | | | | | |
DB 498 CTGCAAGGGGTTTGTCTGGCG 477

RESULT 9

ABL49299/c
ID ABL49299 standard; cDNA; 1743 BP.

XX AC ABL49299;

DT 01-MAY-2002 (first entry)

DE Human lung tumour L523s recombinant polynucleotide SEQ ID NO:450.

KW Human; lung tumour; lung cancer; cytostatic; immunostimulant; vaccine;
KW immune response; ss.

XX OS Homo sapiens.

XX PN WO200200174-A2.

XX PD 03-JAN-2002.

XX PF 28-JUN-2001; 2001WO-US21065.

XX PR 28-JUN-2000; 2000US-0606421.

XX PR 02-AUG-2000; 2000US-0630940.

XX PR 21-AUG-2000; 2000US-0643597.

XX PR 15-SEP-2000; 2000US-0662786.

XX PR 09-OCT-2000; 2000US-0685696.

XX PR 12-DEC-2000; 2000US-0735705.

XX PR 07-MAY-2001; 2001US-0850716.

XX (CORI-) CORIXA CORP.

XX PI Wang T, Wang A, Skeiky YAW, Li SX, Kalos MD, Henderson RA;
PI McNeill PD, Fanger N, Retter MW, Marnerakis M, Fanger GR;
PI Vedvick TS, Carter D, Watanabe Y, Peckham DW;

DR WPI; 2002-090513/12.

XX P-PSDB; ABB75054.

XX PT Polynucleotides encoding lung tumor polypeptides, useful for treating
XX lung cancer or stimulating an immune response -

XX PS Claim 1; Page 370; 374pp; English.

XX CC The present invention describes human lung tumour proteins. Human lung
XX tumour proteins and polynucleotides have cytostatic and immunostimulant
XX activities, and can be used in vaccine production. Compositions
XX comprising the lung tumour proteins, polynucleotides, antibodies,
XX fusion proteins, T cell populations, or antigen presenting cells that
XX express the lung tumour proteins are useful for treating lung cancer or
XX stimulating an immune response. ABL48959 to ABL49300 and ABB74946 to
XX ABB75070 represent sequences used in the exemplification of the present
XX invention.

SQL Sequence 1743 BP; 527 A; 406 C; 418 G; 392 T; 0 other;

Query Match 100.0%; Score 22; DB 24; Length 1743;
Best Local Similarity 100.0%; Pred. No. 0.9;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTGCAAGGGGTTTGTCTGGCG 22
| | | | | | | | | | | | | | | | | | | | | |
DB 498 CTGCAAGGGGTTTGTCTGGCG 477

RESULT 10

ABL49283/c
ID ABL49283 standard; cDNA; 1764 BP.

XX AC ABL49283;

DT 01-MAY-2002 (first entry)

DE Human lung tumour L523s recombinant polynucleotide SEQ ID NO:428.

KW Human; lung tumour; lung cancer; cytostatic; immunostimulant; vaccine;
KW immune response; ss.

XX OS Homo sapiens.

XX PN WO200200174-A2.

XX PD 03-JAN-2002.

XX PF 28-JUN-2001; 2001WO-US21065.

XX PR 28-JUN-2000; 2000US-0606421.

XX PR 02-AUG-2000; 2000US-0630940.

XX PR 21-AUG-2000; 2000US-0643597.

XX PR 15-SEP-2000; 2000US-0662786.

XX PR 09-OCT-2000; 2000US-0685696.

XX PR 12-DEC-2000; 2000US-0735705.

XX PR 07-MAY-2001; 2001US-0850716.

XX (CORI-) CORIXA CORP.

XX PI Wang T, Wang A, Skeiky YAW, Li SX, Kalos MD, Henderson RA;
PI McNeill PD, Fanger N, Retter MW, Marnerakis M, Fanger GR;
PI Vedvick TS, Carter D, Watanabe Y, Peckham DW;

DR WPI; 2002-090513/12.

XX P-PSDB; ABB75048.

XX PT Polynucleotides encoding lung tumor polypeptides, useful for treating
XX lung cancer or stimulating an immune response -

XX PS Claim 1; Page 355-356; 374pp; English.

XX CC The present invention describes human lung tumour proteins. Human lung
XX tumour proteins and polynucleotides have cytostatic and immunostimulant
XX activities, and can be used in vaccine production. Compositions
XX comprising the lung tumour proteins, polynucleotides, antibodies,
XX fusion proteins, T cell populations, or antigen presenting cells that
XX express the lung tumour proteins are useful for treating lung cancer or
XX stimulating an immune response. ABL48959 to ABL49300 and ABB74946 to
XX ABB75070 represent sequences used in the exemplification of the present
XX invention.

SQL Sequence 1764 BP; 534 A; 417 C; 419 G; 394 T; 0 other;

Query Match 100.0%; Score 22; DB 24; Length 1764;
Best Local Similarity 100.0%; Pred. No. 0.9;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTGCAAGGGGTTTGTCTGGCG 22
| | | | | | | | | | | | | | | | | | | | | |
DB 519 CTGCAAGGGGTTTGTCTGGCG 498

RESULT 11

AAZ36150/c
ID AAZ36150 standard; DNA; 4159 BP.

XX AC AAZ36150;

DT 11-FEB-2000 (first entry)

XX DE DNA encoding cancer associated antigen KOC-1.
XX KW Cancer associated antigen; KOC-1; cancer; vaccine; CT7; ss.
XX OS Homo sapiens.
XX PN WO9954738-A1.
XX PD 28-OCT-1999.
XX PF 16-MAR-1999; 99WO-US05766.
XX PR 17-APR-1998; 98US-0061709.
XX PA (LUDW-) LUDWIG INST CANCER RES.
XX PI Chen Y, Gure A, Tsang S, Stockert E, Jager E, Knuth A, Old LJ;
XX DR WPI; 2000-013284/01.
XX Nucleotides representing cancer-associated genes, used to develop
XX PT products for the diagnosis, monitoring and treatment of cancers -
XX PS Claim 88; Page 39-40; 44pp; English.
XX The present sequence represents a cancer associated antigen gene
XX designated KOC-1. The specification also describes a cancer associated
XX antigen designated CT7. The CT7 polynucleotide was isolated from
XX SK-MEL-37 melanoma cells. The polypeptide has some homology with
XX MAGE-10, limited to about 210 carboxy terminal amino acids. The amino
XX terminal of the protein has a repetitive pattern, with repeats rich in
XX serine, proline, glutamine and leucine, and an almost invariable core of
XX the peptide given in AAY43877. The CT7 polypeptide can be processed to
XX peptides which provoke lysis by cytolytic T cells. The polynucleotides
XX and polypeptides can be used for treating a cancerous condition and
XX screening for or diagnosing cancerous conditions. The cancer associated
XX antigens can be used as an immunogenic or vaccine composition with an
XX adjuvant, e.g. a cytokine, a saponin, or granulocyte macrophage-colony
XX stimulating factor (GM-CSF).
XX SQ Sequence 4159 BP; 1281 A; 830 C; 851 G; 1181 T; 16 other;

Query Match 100.0%; Score 22; DB 21; Length 4159;
Best Local Similarity 100.0%; Pred. No. 1;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTGCAAGGGGTTTGTCTGGGCG 22
|||||
Db 748 CTGCAAGGGGTTTGTCTGGGCG 727

RESULT 12
AAC65900/c
ID AAC65900 standard; cDNA; 4181 BP.
XX AC AAC65900;
XX XX
XX 21-FEB-2001 (first entry)
XX XX
XX Human lung cancer-associated cDNA L5238.
XX XX
XX Lung cancer; therapy; treatment; human; tumor; immunogenic; cytostatic;
XX KW vaccine; detection; ss.
XX OS Homo sapiens.
XX OS
XX PN WO200061612-A2.
XX XX
XX PD 19-OCT-2000.
XX XX
XX 03-APR-2000; 2000WO-US08896.
XX XX

PR 02-APR-1999; 99US-0285479.
PR 17-DEC-1999; 99US-0466396.
PR 30-DEC-1999; 99US-0476496.
PR 10-JAN-2000; 2000US-0480884.
PR 22-FEB-2000; 2000US-0510376.
XX (CORI-) CORIXA CORP.
XX XX
XX PI Wang T, Fan L;
XX DR WPI; 2000-628399/60.
XX DR P-PSDB; AAB11328.
XX PT Isolated polypeptide comprising an immunogenic portion of a lung tumor
XX PT protein is used for detecting and monitoring progression of lung cancer
XX PT in a patient -
XX PS Claim 1a; Page 184-186; 261pp; English.
XX This invention describes a novel isolated polypeptide (I) which
XX comprising an immunogenic portion of a lung tumor protein or variant (P2)
XX which have cytostatic activity. The polypeptides and polynucleotides are
XX used in compositions and vaccines to inhibit the development of cancer,
XX especially lung cancer, in a patient. Methods described in the invention
XX can be used to monitor the progression of a cancer by carrying out the
XX detection at subsequent time points and comparing the results from the
XX different time points. CD4+ and/or CD8+ T-Cells isolated from a patient
XX are treated with P2, polynucleotides encoding P2 or antigen presenting
XX cells expressing P2 and then administered to the patient to inhibit
XX development of cancer.
XX SQ Sequence 4181 BP; 1303 A; 830 C; 851 G; 1181 T; 16 other;

Query Match 100.0%; Score 22; DB 21; Length 4181;
Best Local Similarity 100.0%; Pred. No. 1;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTGCAAGGGGTTTGTCTGGGCG 22
|||||
Db 748 CTGCAAGGGGTTTGTCTGGGCG 727

RESULT 13
ABL65405/c
ID ABL65405 standard; DNA; 4181 BP.
XX AC ABL65405;
XX XX
XX 15-MAY-2002 (first entry)
XX XX
XX Lung cancer related gene sequence SEQ ID NO:3742.
XX XX
XX Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;
XX KW stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;
XX KW cytostatic; gene therapy; antineoplastic; Wilm's tumour; adenocarcinoma;
XX KW gene; ds.
XX OS Homo sapiens.
XX XX
XX WO200194629-A2.
XX XX
XX 13-DEC-2001.
XX XX
XX 30-MAY-2001; 2001WO-US10838.
XX XX
XX 05-JUN-2000; 2000US-209473P.
XX PR 05-JUN-2000; 2000US-209531P.
XX PR 18-SEP-2000; 2000US-233133P.
XX PR 18-SEP-2000; 2000US-233617P.
XX PR 20-SEP-2000; 2000US-234009P.
XX PR 20-SEP-2000; 2000US-234034P.
XX PR 20-SEP-2000; 2000US-234052P.
XX PR 22-SEP-2000; 2000US-234509P.

PR 22-SEP-2000; 2000US-234567P.
PR 25-SEP-2000; 2000US-234923P.
PR 25-SEP-2000; 2000US-234924P.
PR 25-SEP-2000; 2000US-235077P.
PR 25-SEP-2000; 2000US-235082P.
PR 25-SEP-2000; 2000US-235134P.
PR 25-SEP-2000; 2000US-235280P.
PR 26-SEP-2000; 2000US-235637P.
PR 27-SEP-2000; 2000US-235638P.
PR 27-SEP-2000; 2000US-235711P.
PR 27-SEP-2000; 2000US-235720P.
PR 27-SEP-2000; 2000US-235840P.
PR 27-SEP-2000; 2000US-235863P.
PR 28-SEP-2000; 2000US-236028P.
PR 28-SEP-2000; 2000US-236032P.
PR 28-SEP-2000; 2000US-236033P.
PR 28-SEP-2000; 2000US-236034P.
PR 28-SEP-2000; 2000US-236109P.
PR 28-SEP-2000; 2000US-236110P.
PR 29-SEP-2000; 2000US-236842P.
PR 29-SEP-2000; 2000US-236891P.
PR 02-OCT-2000; 2000US-237172P.
PR 02-OCT-2000; 2000US-237173P.
PR 02-OCT-2000; 2000US-237278P.
PR 02-OCT-2000; 2000US-237294P.
PR 02-OCT-2000; 2000US-237295P.
PR 02-OCT-2000; 2000US-237316P.
PR 03-OCT-2000; 2000US-237425P.
PR 03-OCT-2000; 2000US-237598P.
PR 03-OCT-2000; 2000US-237604P.
PR 03-OCT-2000; 2000US-237606P.
PR 03-OCT-2000; 2000US-237608P.
PR 01-NOV-2000; 2000US-244867P.
PR 01-NOV-2000; 2000US-245084P.
XX

(AVAL-) AVALON PHARM.

XX Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;
XX Soppet DR, Weaver Z;
XX WPI; 2002-188264/24.

PT Screening for anti-neoplastic agent involves exposing cells to a
PT chemical agent to be tested for anti-neoplastic activity, and
PT determining a change in expression of a gene of a signature gene set

XX Claim 1; SEQ ID 3742; 44pp; English.

XX The present invention describes a method (M1) for screening for an
CC anti-neoplastic agent. The method involves exposing cells to a chemical
CC agent to be tested for anti-neoplastic activity, determining a change in
CC expression of at least one gene (I) of a signature gene set, where (I)
CC comprises a sequence (S) selected from 8447 sequences (given in ABL61664
CC to ABL70110), or is at least 95% identical to (S), where a change in
CC expression is indicative of anti-neoplastic activity. (I) has cytostatic
CC activity and can be used in gene therapy. M1 can be used for screening
CC an anti-neoplastic agent, and can be used for producing a product which
CC is the data collected with respect to the anti-neoplastic agent as a
CC result of M1, and the data is sufficient to convey the chemical
CC structure and/or properties of the agent. M1 can be used in the
CC treatment of cancer such as colon, breast, stomach, lung, thyroid,
CC oesophageal, ovarian, kidney, prostate or pancreatic cancer.
CC adenocarcinoma, carcinoma, clear cell cancer, infiltrating ductal cancer,
CC infiltrating lobular cancer, squamous cell carcinoma, neuroendocrine
CC carcinoma, papillary carcinoma and Wilms' tumour.
XX

XX Sequence 4181 BP; 1303 A; 830 C; 851 G; 1181 T; 16 other;

Query Match 100.0%; Score 22; DB 24; Length 4181;
Best Local Similarity 100.0%; Pred. No. 1;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTGCAAGGGGTTTGTCTGGCG 22

Db 748 CTGCAAGGGGTTTGTCTGGCG 727

RESULT 14

ABL49119/c
ID ABL49119 standard; cDNA; 4181 BP.

XX ABL49119;

XX 01-MAY-2002 (first entry)

XX Human lung tumour LS23S cDNA sequence SEQ ID NO:175.

XX Human; lung tumour; lung cancer; cytostatic; immunostimulant; vaccine;
KW immune response; ss.

XX Homo sapiens.

XX WO200200174-A2.

XX 03-JAN-2002.

XX 28-JUN-2001; 2001WO-US21065.

XX 28-JUN-2000; 2000US-0606421.

XX 02-AUG-2000; 2000US-0630940.

XX 21-AUG-2000; 2000US-0643597.

XX 15-SEP-2000; 2000US-0662786.

XX 09-OCT-2000; 2000US-0685696.

XX 12-DEC-2000; 2000US-0735705.

XX 07-MAY-2001; 2001US-0850716.

XX (CORI-) CORIXA CORP.

XX Wang T, Wang A, Skeiky YAW, Li SX, Kalos MD, Henderson RA;
XX McNeill PD, Fanger N, Retter MW, Marnerakis M, Fanger GR;
XX Vedvick TS, Carter D, Watanabe Y, Peckham DW;

XX WPI; 2002-090513/12.

XX P-PSDB; ABB74960.

XX Polynucleotides encoding lung tumor polypeptides, useful for treating
XX lung cancer or stimulating an immune response -

XX Example 2; Page 266-267; 374pp; English.

XX The present invention describes human lung tumour proteins. Human lung
CC tumour proteins and polynucleotides have cytostatic and immunostimulant
CC activities, and can be used in vaccine production. Compositions
CC comprising the lung tumour proteins, polynucleotides, antibodies,
CC fusion proteins, T cell populations, or antigen presenting cells that
CC express the lung tumour proteins are useful for treating lung cancer or
CC stimulating an immune response. ABL4959 to ABL49300 and ABB74946 to
CC ABB75070 represent sequences used in the exemplification of the present
CC invention.

XX Sequence 4181 BP; 1303 A; 830 C; 851 G; 1181 T; 16 other;

Query Match 100.0%; Score 22; DB 24; Length 4181;
Best Local Similarity 100.0%; Pred. No. 1;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTGCAAGGGGTTTGTCTGGCG 22

Db 748 CTGCAAGGGGTTTGTCTGGCG 727

RESULT 15

ABA99958/c

ID ABA99958 standard; DNA; 4601 BP.

XX ABA99958;

Mon Apr 21 10:29:40 2003

Search completed: April 18, 2003, 05:45:18
Job time : 100 secs

XX 05-JUL-2002 (first entry)
XX Human KOC DNA.
XX KOC; promoter; human; cytostatic; dermatological; vulnary; tumour;
XX Chemotherapy; allogenic bone marrow transplant; pancreatic carcinoma;
XX domain-containing protein overexpressed in cancer; immunostimulant;
XX gene therapy; chronic pancreatitis; pluripotency; aging; wound healing;
XX radiation therapy; skin regeneration; cell differentiation; cancer; ds.
XX Homo sapiens.
XX WO200220036-A1.
XX 14-MAR-2002.
XX 01-AUG-2001; 2001WO-DE02948.
XX 06-SEP-2000; 2000DE-1043964.
XX 14-FEB-2001; 2001DE-1006829.
XX (MUEL/) MUELLER F.
XX Mueller F, Gress T, Adler G;
XX WPI; 2002-292230/33.
XX New promoter of the KOC gene, useful for diagnosis and treatment of
XX cancer, for inducing cellular dedifferentiation and for identifying
XX modulators
XX Claim 2; Fig 1a-b; 74pp; German.
XX This invention describes a novel KOC (KH domain-containing protein
XX overexpressed in cancer) promoter. The products of the invention have
XX cytostatic, dermatological, vulnary and immunostimulant activity, are
XX capable of modulating activity/expression of the KOC protein and can be
XX used for gene therapy. The KOC promoter is also useful (i) for diagnosis
XX of tumours that express KOC, particularly for differentiating between
XX chronic pancreatitis and pancreatic carcinoma; (ii) to detect
XX premalignant lesions of uncertain nature and to classify them for risk,
XX and (iii) to evaluate treatments; (iv) to impart pluripotency to cells;
XX (v) to prepare tissues or organs (by differentiation of stem cell
XX populations); (vi) in high-dose chemotherapy; (vii) to improve ex vivo
XX expansion of hematopoietic stem cells; (viii) to slow down or reverse aging
XX allogenic bone marrow transplants; (ix) to improve engraftment of
XX processes; (x) as prophylactic agents during chemotherapy/radiation
XX therapy; (xi) to regenerate skin defects and accelerate wound healing,
XX including for cosmetic purposes; and (xii) to immunise against malignant
XX tumours or their precursor stages. The promoter can also be used (i) for
XX isolation/enrichment/selective replication of stem cells; (ii) to
XX determine the degree of (de)differentiation of cells and tissues; (iii)
XX in gene therapy of cancer (for controlling protein expression); and (iv)
XX to identify compounds, potentially useful in cancer treatment, that bind
XX to the promoter and modulate its activity. The promoter makes possible
XX generation of stem cells of any selected type. Since expression of KOC
XX is not sensitive to other physiological states, e.g. inflammation, it
XX represents a tumour marker with 100% sensitivity and specificity. This
XX sequence represents the human KOC gene described in the disclosure of
XX the invention.
XX SQ Sequence 4601 BP; 1462 A; 906 C; 918 G; 1298 T; 17 other;
Query Match 100.0%; Score 22; DB 24; Length 4601;
Best Local Similarity 100.0%; Pred. No. 1;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CTGCAAGGGGTTTCTGCGGCG 22
DB 1169 CTGCAAGGGGTTTCTGCGGCG 1147

Result No.	Score	Query		Length	DB	ID	Description
		Match	%				
C 1	22	100.0	1740	4	US-09-643-597-347		Sequence 347, Appl
C 2	22	100.0	4159	4	US-09-061-709-4		Sequence 4, Appli
C 3	22	100.0	4181	4	US-09-643-597-175		Sequence 175, App
C 4	16.2	73.6	5552	4	US-09-676-974-2		Sequence 2, Appli
C 5	15.6	70.9	1863	3	US-09-199-229-1		Sequence 1, Appli
C 6	15.6	70.9	1863	3	US-09-443-087-1		Sequence 1, Appli
C 7	15.6	70.9	1863	3	US-09-687-298-1		Sequence 1, Appli
C 8	15.6	70.9	3095	4	US-09-325-932A-128		Sequence 128, App
C 9	15.6	70.9	3487	4	US-09-453-702B-164		Sequence 164, App
C 10	15.6	70.9	11827	4	US-09-739-455-3		Sequence 3, Appli
C 11	15.6	70.9	48908	4	US-09-453-702B-137		Sequence 137, App
C 12	15.4	70.0	5436	2	US-08-948-277A-2		Sequence 2, Appli
C 13	15.4	70.0	5436	2	US-09-169-203-2		Sequence 2, Appli
C 14	15.2	69.1	726	4	US-09-830-118-151		Sequence 151, App
C 15	15.2	69.1	857	4	US-08-282-488-1		Sequence 1, Appli
C 16	15.2	69.1	865	3	US-09-040-483-2		Sequence 2, Appli
C 17	15.2	69.1	4208	4	US-09-214-278-6		Sequence 6, Appli
C 18	15.2	69.1	4208	4	US-09-068-740A-10		Sequence 10, Appli
C 19	15.2	69.1	5029	4	US-09-462-138-3		Sequence 3, Appli
C 20	15.2	69.1	5458	4	US-09-199-865-2		Sequence 2, Appli
C 21	15.2	69.1	5590	3	US-08-882-046-1		Sequence 1, Appli
C 22	15.2	69.1	6464	2	US-08-400-159-5		Sequence 5, Appli
C 23	15.2	69.1	6464	3	US-08-611-729A-5		Sequence 5, Appli
C 24	15.2	69.1	7130	4	US-09-056-105-31		Sequence 31, Appli
C 25	15	68.2	4368	1	US-08-026-138E-17		Sequence 17, Appli
C 26	15	68.2	4446	1	US-08-026-138E-6		Sequence 6, Appli
C 27	14.8	67.3	908	4	US-09-605-785-350		Sequence 350, App

;; TITLE OF INVENTION: Isolated Nucleic Acid Molecules Encoding Cancer Associated
;; FILE OF INVENTION: Antigen, The Antigens Per Se, And Uses Thereof
;; FILE REFERENCE: LUD 5538
;; CURRENT APPLICATION NUMBER: US/09/061,709B
;; CURRENT FILING DATE: 1998-04-17
;; NUMBER OF SEQ ID NOS: 8
;; SEQ ID NO 4

;; LENGTH: 4159
;; TYPE: DNA
;; ORGANISM: Homo sapiens
;; FEATURE:
US-09-061-709-4

Query Match 100.0%; Score 22; DB 4; Length 4159;
Best Local Similarity 100.0%; Pred. No. 0.091;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTGCAAGGGGTTTGTCTGGCG 22
|||||
DB 748 CTGCAAGGGGTTTGTCTGGCG 727

RESULT 3

US-09-643-597-175/c
; Sequence 175, Application US/09643597
; Patent No. 6426072

GENERAL INFORMATION:

;; APPLICANT: Wang, Tongtong
;; APPLICANT: Fan, Liqun
;; APPLICANT: Kalos, Michael D.
;; APPLICANT: Bangur, Chaitanya S.
;; APPLICANT: Hosken, Nancy
;; APPLICANT: Fanger, Gary R.
;; APPLICANT: Li, Samuel X.
;; APPLICANT: Wang, Aijun
;; APPLICANT: Skeiky, Yasir A.W.
;; APPLICANT: Henderson, Robert A.
;; APPLICANT: McNeill, Patricia D.

;; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
;; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER

;; FILE REFERENCE: 210121.455C11
;; CURRENT APPLICATION NUMBER: US/09/643,597

;; CURRENT FILING DATE: 2000-08-21
;; NUMBER OF SEQ ID NOS: 369

;; SOFTWARE: FastSeq for Windows Version 3.0
;; SEQ ID NO 175

;; LENGTH: 4181

;; TYPE: DNA

;; ORGANISM: Homo sapiens

;; FEATURE:

;; NAME/KEY: unsure

;; LOCATION: (3347)

;; OTHER INFORMATION: n=A,T,C or G

;; NAME/KEY: unsure

;; LOCATION: (3502)

;; OTHER INFORMATION: n=A,T,C or G

;; NAME/KEY: unsure

;; LOCATION: (3506)

;; OTHER INFORMATION: n=A,T,C or G

;; NAME/KEY: unsure

;; LOCATION: (3520)

;; OTHER INFORMATION: n=A,T,C or G

;; NAME/KEY: unsure

;; LOCATION: (3538)

;; OTHER INFORMATION: n=A,T,C or G

;; NAME/KEY: unsure

;; LOCATION: (3549)

;; OTHER INFORMATION: n=A,T,C or G

;; NAME/KEY: unsure

;; LOCATION: (3646)

;; OTHER INFORMATION: n=A,T,C or G

;; NAME/KEY: unsure

;; LOCATION: (3940)

;; OTHER INFORMATION: n=A,T,C or G
;; NAME/KEY: unsure
;; LOCATION: (3968)
;; OTHER INFORMATION: n=A,T,C or G
;; NAME/KEY: unsure
;; LOCATION: (3974)

;; OTHER INFORMATION: n=A,T,C or G
;; NAME/KEY: unsure
;; LOCATION: (4036)

;; OTHER INFORMATION: n=A,T,C or G
;; NAME/KEY: unsure
;; LOCATION: (4056)

;; OTHER INFORMATION: n=A,T,C or G
;; NAME/KEY: unsure
;; LOCATION: (4062)

;; OTHER INFORMATION: n=A,T,C or G
;; NAME/KEY: unsure
;; LOCATION: (4088)

;; OTHER INFORMATION: n=A,T,C or G
;; NAME/KEY: unsure
;; LOCATION: (4115)

;; OTHER INFORMATION: n=A,T,C or G
;; NAME/KEY: unsure
;; LOCATION: (4115)

;; OTHER INFORMATION: n=A,T,C or G
;; NAME/KEY: unsure
;; LOCATION: (4115)

;; OTHER INFORMATION: n=A,T,C or G
;; NAME/KEY: unsure
;; LOCATION: (4115)

;; OTHER INFORMATION: n=A,T,C or G
;; NAME/KEY: unsure
;; LOCATION: (4115)

;; OTHER INFORMATION: n=A,T,C or G
;; NAME/KEY: unsure
;; LOCATION: (4115)

;; OTHER INFORMATION: n=A,T,C or G
;; NAME/KEY: unsure
;; LOCATION: (4115)

;; OTHER INFORMATION: n=A,T,C or G
;; NAME/KEY: unsure
;; LOCATION: (4115)

;; OTHER INFORMATION: n=A,T,C or G
;; NAME/KEY: unsure
;; LOCATION: (4115)

;; OTHER INFORMATION: n=A,T,C or G
;; NAME/KEY: unsure
;; LOCATION: (4115)

;; OTHER INFORMATION: n=A,T,C or G
;; NAME/KEY: unsure
;; LOCATION: (4115)

;; OTHER INFORMATION: n=A,T,C or G
;; NAME/KEY: unsure
;; LOCATION: (4115)

;; OTHER INFORMATION: n=A,T,C or G
;; NAME/KEY: unsure
;; LOCATION: (4115)

;; OTHER INFORMATION: n=A,T,C or G
;; NAME/KEY: unsure
;; LOCATION: (4115)

;; OTHER INFORMATION: n=A,T,C or G
;; NAME/KEY: unsure
;; LOCATION: (4115)

;; OTHER INFORMATION: n=A,T,C or G
;; NAME/KEY: unsure
;; LOCATION: (4115)

;; OTHER INFORMATION: n=A,T,C or G
;; NAME/KEY: unsure
;; LOCATION: (4115)

;; OTHER INFORMATION: n=A,T,C or G
;; NAME/KEY: unsure
;; LOCATION: (4115)

;; OTHER INFORMATION: n=A,T,C or G
;; NAME/KEY: unsure
;; LOCATION: (4115)

;; OTHER INFORMATION: n=A,T,C or G
;; NAME/KEY: unsure
;; LOCATION: (4115)

;; OTHER INFORMATION: n=A,T,C or G
;; NAME/KEY: unsure
;; LOCATION: (4115)

;; OTHER INFORMATION: n=A,T,C or G
;; NAME/KEY: unsure
;; LOCATION: (4115)

;; OTHER INFORMATION: n=A,T,C or G
;; NAME/KEY: unsure
;; LOCATION: (4115)

;; OTHER INFORMATION: n=A,T,C or G
;; NAME/KEY: unsure
;; LOCATION: (4115)

;; OTHER INFORMATION: n=A,T,C or G
;; NAME/KEY: unsure
;; LOCATION: (4115)

;; OTHER INFORMATION: n=A,T,C or G
;; NAME/KEY: unsure
;; LOCATION: (4115)

;; OTHER INFORMATION: n=A,T,C or G
;; NAME/KEY: unsure
;; LOCATION: (4115)

;; OTHER INFORMATION: n=A,T,C or G
;; NAME/KEY: unsure
;; LOCATION: (4115)

;; OTHER INFORMATION: n=A,T,C or G
;; NAME/KEY: unsure
;; LOCATION: (4115)

;; OTHER INFORMATION: n=A,T,C or G
;; NAME/KEY: unsure
;; LOCATION: (4115)

;; OTHER INFORMATION: n=A,T,C or G
;; NAME/KEY: unsure
;; LOCATION: (4115)

;; OTHER INFORMATION: n=A,T,C or G
;; NAME/KEY: unsure
;; LOCATION: (4115)

;; OTHER INFORMATION: n=A,T,C or G
;; NAME/KEY: unsure
;; LOCATION: (4115)

;; OTHER INFORMATION: n=A,T,C or G
;; NAME/KEY: unsure
;; LOCATION: (4115)

;; OTHER INFORMATION: n=A,T,C or G
;; NAME/KEY: unsure
;; LOCATION: (4115)

;; OTHER INFORMATION: n=A,T,C or G
;; NAME/KEY: unsure
;; LOCATION: (4115)

;; OTHER INFORMATION: n=A,T,C or G
;; NAME/KEY: unsure
;; LOCATION: (4115)

;; OTHER INFORMATION: n=A,T,C or G
;; NAME/KEY: unsure
;; LOCATION: (4115)

;; OTHER INFORMATION: n=A,T,C or G
;; NAME/KEY: unsure
;; LOCATION: (4115)

;; OTHER INFORMATION: n=A,T,C or G
;; NAME/KEY: unsure
;; LOCATION: (4115)

;; OTHER INFORMATION: n=A,T,C or G
;; NAME/KEY: unsure
;; LOCATION: (4115)

;; OTHER INFORMATION: n=A,T,C or G
;; NAME/KEY: unsure
;; LOCATION: (4115)

;; OTHER INFORMATION: n=A,T,C or G
;; NAME/KEY: unsure
;; LOCATION: (4115)

;; OTHER INFORMATION: n=A,T,C or G
;; NAME/KEY: unsure
;; LOCATION: (4115)

;; OTHER INFORMATION: n=A,T,C or G
;; NAME/KEY: unsure
;; LOCATION: (4115)

;; OTHER INFORMATION: n=A,T,C or G
;; NAME/KEY: unsure
;; LOCATION: (4115)

;; OTHER INFORMATION: n=A,T,C or G
;; NAME/KEY: unsure
;; LOCATION: (4115)

;; OTHER INFORMATION: n=A,T,C or G
;; NAME/KEY: unsure
;; LOCATION: (4115)

;; OTHER INFORMATION: n=A,T,C or G
;; NAME/KEY: unsure
;; LOCATION: (4115)

;; OTHER INFORMATION: n=A,T,C or G
;; NAME/KEY: unsure
;; LOCATION: (4115)

;; OTHER INFORMATION: n=A,T,C or G
;; NAME/KEY: unsure
;; LOCATION: (4115)

;; OTHER INFORMATION: n=A,T,C or G
;; NAME/KEY: unsure
;; LOCATION: (4115)

;; OTHER INFORMATION: n=A,T,C or G
;; NAME/KEY: unsure
;; LOCATION: (4115)

;; OTHER INFORMATION: n=A,T,C or G
;; NAME/KEY: unsure
;; LOCATION: (4115)

;; OTHER INFORMATION: n=A,T,C or G
;; NAME/KEY: unsure
;; LOCATION: (4115)

;; OTHER INFORMATION: n=A,T,C or G
;; NAME/KEY: unsure
;; LOCATION: (4115)

;; OTHER INFORMATION: n=A,T,C or G
;; NAME/KEY: unsure
;; LOCATION: (4115)

;; OTHER INFORMATION: n=A,T,C or G
;; NAME/KEY: unsure
;; LOCATION: (4115)

;; OTHER INFORMATION: n=A,T,C or G
;; NAME/KEY: unsure
;; LOCATION: (4115)

;; OTHER INFORMATION: n=A,T,C or G
;; NAME/KEY: unsure
;; LOCATION: (4115)

;; OTHER INFORMATION: n=A,T,C or G
;; NAME/KEY: unsure
;; LOCATION: (4115)

;; OTHER INFORMATION: n=A,T,C or G
;; NAME/KEY: unsure
;; LOCATION: (4115)

;; OTHER INFORMATION: n=A,T,C or G
;; NAME/KEY: unsure
;; LOCATION: (4115)

;; OTHER INFORMATION: n=A,T,C or G
;; NAME/KEY: unsure
;; LOCATION: (4115)

;; OTHER INFORMATION: n=A,T,C or G
;; NAME/KEY: unsure
;; LOCATION: (4115)

```

; OTHER INFORMATION: A, T, C or G
; NAME/KEY: modified_base
; LOCATION: (5400)
; OTHER INFORMATION: A, T, C or G
US-09-676-974-2

```

Query Match 73.6%; Score 16.2; DB 4; Length 5552;
Best Local Similarity 85.7%; Pred. No. 63;
Matches 18; Conservative 0; Mismatches 3; Indels 0

Qy 2 TGCAAGGGGTTTTCCTGGGCG 22
Db 1798 TGCAAGGGGTTTTCGGGGGCG 1818

RESULT 5
US-09-199-229-1/c
; Sequence 1, Application US/09199229
; Patent No. 6063607

```

; GENERAL INFORMATION:
; APPLICANT: Yaver, Debbie S.
; APPLICANT: Berkta, Randy M.
; APPLICANT: Rev. Michael W.
; TITLE OF INVENTION: Polypeptides Having Choline Oxidase
; FILE OF INVENTION: Activity and Nucleic Acids Encoding Same
; FILE REFERENCE: 5735.000-US
; CURRENT APPLICATION NUMBER: US/09/199,229
; CURRENT FILING DATE: 1998-11-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1

```

```

; LENGTH: 1863
; TYPE: DNA
; ORGANISM: Fusarium
US-09-199-229-1

```

Query Match 70.9%; Score 15.6; DB 3; Length 1863;
Best Local Similarity 81.8%; Pred.No. 99;
Matches 18; Conservative 0; Mismatches 4; Indels 0

Qy 1 CTGCAAGGGGTTTTCGTGGCG 22
pb 1359 CTTGAAGGGGCTCTGCTGGCG 1338

RESULT 6
US-09-443-087-1/c
; Sequence 1, Application US/09443087
; Patent NO. 6146864

/ GENERAL INFORMATION:
 / APPLICANT: Debbie Yaver
 / APPLICANT: Randy M. Berka
 / APPLICANT: Michael W. Rey
 / TITLE OF INVENTION: Polypeptides Having Choline Oxidase
 / TITLE OF INVENTION: Activity And Nucleic Acids Encoding Same
 / FILE REFERENCE: 5735,200-US
 / CURRENT APPLICATION NUMBER: US/09/443,087
 / CURRENT FILING DATE: 1999-11-18
 / EARLIER APPLICATION NUMBER: 09/199,229
 / EARLIER FILING DATE: 1998-11-24
 / NUMBER OF SEQ ID NOS: 2

```

; SOFTWARE: F8B
; SEQ ID NO 1
; LENGTH: 1863
; TYPE: DNA

```

US-09-443-087-1

Query Match 70.9%; Score 15.6; DB 3; Length 1863;
Best Local Similarity 81.8%; Pred. No. 99;
Matches 18; Conservative 0; Mismatches 4; Indels 0

Qy 1 CTGCAAGGGGTTTGTCTGGGCG 22

Db 1359 CTTGAAGGGGCTCTGCTGGCG 1338

RESULT 7
US-09-687-298-1/c
; Sequence 1, Application US/09687298

BEST LOCAL SIMILARITY 85.7%; FREQ. NO. 63;
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```

, APPLICANT: Randy M. Berka
, APPLICANT: Michael W. Rey
, TITLE OF INVENTION: Polypeptides Having Choline Oxidase
, TITLE OF INVENTION: Activity And Nucleic Acids Encoding Same
, FILE REFERENCE: 5735,200-US
, CURRENT APPLICATION NUMBER: US/09/687,298
, CURRENT FILING DATE: 2000-10-12
, PRIOR APPLICATION NUMBER: 09/443,087
, PRIOR FILING DATE: 1999-11-18
, NUMBER OF SEQ ID NOS: 2
, SOFTWARE: FastSeq for Windows Version 3.0
, SEQ ID NO 1
, LENGTH: 1863
, TYPE: DNA
, ORGANISM: Fusarium venenatum
, US-09-687-298-1

```

Query Match 70.9%; Score 15.6; DB 4; Length 1863;
Best Local Similarity 81.8%; Pred. No. 99;
Matches 18: Conservative 0; Mismatches 4; Indels 0

Qy 1 CTGCAAGGGTTTTGCTGGCG 22
Db 1359 CTTGAAGGGGCTCTGCTGGCG 1338

RESULT 8
US-09-325-932A-128/c
; Sequence 128, Application US/09325932A
; Patent No. 6451604

```

/ GENERAL INFORMATION:
/ APPLICANT: Flinn, Barry
/ APPLICANT: Lashau, Annette
/ TITLE OF INVENTION: Compositions affecting programmed cell
/ : death and their use in the modification
/ TITLE OF INVENTION:
/ FILE REFERENCE: 1022
/ CURRENT APPLICATION NUMBER: US/09/325,932A
/ CURRENT FILING DATE: 1999-06-04
/ NUMBER OF SEQ ID NOS: 206
/ SOFTWARE: FastSeq for Windows Version 3.0
/ SEQ ID NO 128
/ LENGTH: 3095
/ TYPE: DNA
/ ORGANISM: Pinus radiata
/ US-09-325-932A-128

```

Query Match 70.9%; Score 15.6; DB 4; Length 3095;
Best Local Similarity 81.8%; Pred. No. 1.1e+02;
Matches 18: Conservative 0; Mismatches 4; Indels 0

Qy 1 CTGCAAGGGGTTTGTCTGGCG 22
||| ||||||||| |||||
Dh 2984 CTCCTCAGGGGTTTTTCAGGGCG 2963

RESULT 9
US-09-453-702B-164/c
; Sequence 164, Application US/09453702B
; Patent No. 6365723
; GENERAL INFORMATION:
; APPLICANT: Blattner, Frederick R.
; Burland, Valerie
; Perna, Nicole T.
;

Plunkett, Guy
Welch, Rod
TITLE OF INVENTION: No. 6365723el Sequences of E. coli O157
NUMBER OF SEQUENCES: 265
CORRESPONDENCE ADDRESS:
ADDRESSEE: Quarles & Brady
STREET: 1 South Pinckney Street
CITY: Madison
STATE: WI
COUNTRY: US
ZIP: 53701-2113
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch. 1.44Mb storage
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Word Perfect 8.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/453,702B
FILING DATE: 03-Dec-1999
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/110,955
FILING DATE: 04-DEC-1998
ATTORNEY/AGENT INFORMATION:
NAME: Seay, Nicholas J.
REGISTRATION NUMBER: 27386
REFERENCE/DOCKET NUMBER: 960296.95017
TELECOMMUNICATION INFORMATION:
TELEPHONE: (608) 251-5000
TELEFAX: (608) 251-9166
INFORMATION FOR SEQ ID NO: 164:
SEQUENCE CHARACTERISTICS:
LENGTH: 3487
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
SEQUENCE DESCRIPTION: SEQ ID NO: 164:
US-09-453-702B-164

Query Match 70.9%; Score 15.6; DB 4; Length 3487;
Best Local Similarity 81.8%; Pred. No. 1.1e+02;
Matches 18; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 CTGCAAGGGGTTTGTCTGGCG 22
Db 1644 CTGGAAGTGGTTTGTCTGCTCG 1623

RESULT 10
US-09-739-455-3/c
Sequence 3, Application US/09739455
Patent No. 6413756
GENERAL INFORMATION:
APPLICANT: YAN, Chunhua et al
TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
TITLE OF INVENTION: THEREOF
FILE REFERENCE: CL000653
CURRENT APPLICATION NUMBER: US/09/739,455
CURRENT FILING DATE: 2000-12-19
NUMBER OF SEQ ID NOS: 23
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 3
LENGTH: 11827
TYPE: DNA
ORGANISM: Human
FEATURE:
NAME/KEY: misc feature
LOCATION: (1)...(11827)
OTHER INFORMATION: n = A,T,C or G
US-09-739-455-3

Query Match 70.9%; Score 15.6; DB 4; Length 11827;
Best Local Similarity 81.8%; Pred. No. 1.4e+02;
Matches 18; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
Qy 1 CTGCAAGGGGTTTGTCTGGCG 22
Db 10677 CTGCTGGGGTGTCTGTGAGCG 10656

RESULT 11
US-09-453-702B-137
Sequence 137, Application US/09453702B
Patent No. 6365723
GENERAL INFORMATION:
APPLICANT: Blattner, Frederick R.
Burland, Valerie
Perna, Nicole T.
Plunkett, Guy
Welch, Rod
TITLE OF INVENTION: No. 6365723el Sequences of E. coli O157
NUMBER OF SEQUENCES: 265
CORRESPONDENCE ADDRESS:
ADDRESSEE: Quarles & Brady
STREET: 1 South Pinckney Street
CITY: Madison
STATE: WI
COUNTRY: US
ZIP: 53701-2113
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch. 1.44Mb storage
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Word Perfect 8.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/453,702B
FILING DATE: 03-Dec-1999
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/110,955
FILING DATE: 04-DEC-1998
ATTORNEY/AGENT INFORMATION:
NAME: Seay, Nicholas J.
REGISTRATION NUMBER: 27386
REFERENCE/DOCKET NUMBER: 960296.95017
TELECOMMUNICATION INFORMATION:
TELEPHONE: (608) 251-5000
TELEFAX: (608) 251-9166
INFORMATION FOR SEQ ID NO: 137:
SEQUENCE CHARACTERISTICS:
LENGTH: 48908
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
SEQUENCE DESCRIPTION: SEQ ID NO: 137:
US-09-453-702B-137

Query Match 70.9%; Score 15.6; DB 4; Length 48908;
Best Local Similarity 81.8%; Pred. No. 1.8e+02;
Matches 18; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 CTGCAAGGGGTTTGTCTGGCG 22
Db 18224 CTGGATGGGTTTGTCTGGCG 18245

RESULT 12
US-08-948-277A-2
Sequence 2, Application US/08948277A
Patent No. 5849581
GENERAL INFORMATION:
APPLICANT: Anaral, M. Catherine
Zhang, Ning

APPLICANT: Chen, Jin-Long
TITLE OF INVENTION: Regulators of UCP3 Gene Expression
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
STREET: 75 DENISE DRIVE
CITY: HILLSBOROUGH
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94010
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/948,277A
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: OSMAN, RICHARD A
REGISTRATION NUMBER: 36,627
REFERENCE/DOCKET NUMBER: T97-009
TELECOMMUNICATION INFORMATION:
TELEPHONE: (650) 343-4341
TELEFAX: (650) 343-4342
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 5436 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-948-277A-2
Query Match 70.0%; Score 15.4; DB 2; Length 5436;
Best Local Similarity 94.1%; Pred. No. 1.5e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 CTGCAAGGGGTTTGCT 17
|||||
DB 4007 CTGCAAGGGGATTGCT 4023
RESULT 13
US-09-169-203-2
Sequence 2, Application US/09169203
Patent No. 5976808
GENERAL INFORMATION:
APPLICANT: Amaral, M. Catherine
APPLICANT: Zhang, Ning
APPLICANT: Chen, Jin-Long
TITLE OF INVENTION: Regulators of UCP3 Gene Expression
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
STREET: 75 DENISE DRIVE
CITY: HILLSBOROUGH
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94010
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/169,203
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/948,277
FILING DATE:

ATTORNEY/AGENT INFORMATION:
NAME: OSMAN, RICHARD A
REGISTRATION NUMBER: 36,627
REFERENCE/DOCKET NUMBER: T97-009
TELECOMMUNICATION INFORMATION:
TELEPHONE: (650) 343-4341
TELEFAX: (650) 343-4342
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 5436 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-09-169-203-2
Query Match 70.0%; Score 15.4; DB 2; Length 5436;
Best Local Similarity 94.1%; Pred. No. 1.5e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 CTGCAAGGGGTTTGCT 17
|||||
DB 4007 CTGCAAGGGGATTGCT 4023
RESULT 14
US-09-280-116-151/c
Sequence 151, Application US/09280116A
Patent No. 6331427
GENERAL INFORMATION:
APPLICANT: Robison, Keith E.
TITLE OF INVENTION: Nucleic Acid Molecules Encoding Human Protease Homologs
FILE REFERENCE: 5800-24, 035800/176965
CURRENT APPLICATION NUMBER: US/09/280,116A
CURRENT FILING DATE: 1999-03-26
NUMBER OF SEQ ID NOS: 268
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 151
LENGTH: 726
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: trypsin-like serine proteases
FEATURE:
NAME/KEY: misc feature
LOCATION: (1)-(726)
OTHER INFORMATION: n = a, t, c or g
US-09-280-116-151
Query Match 69.1%; Score 15.2; DB 4; Length 726;
Best Local Similarity 85.0%; Pred. No. 1.3e+02;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 2 TCGAAGGGGTTTGCTGGC 21
|||||
DB 615 TCGAAGGGGTTTGCAAGCC 596
RESULT 15
US-08-832-488-1
Sequence 1, Application US/08832488
Patent No. 6448044
GENERAL INFORMATION:
APPLICANT: NI, JIAN
APPLICANT: LI, HAODONG
APPLICANT: YU, GUO-LIANG
APPLICANT: GENTZ, REINER L.
TITLE OF INVENTION: HUMAN NATURAL KILLER CELL ACTIVATION
TITLE OF INVENTION: FACTOR II
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: HUMAN GENOME SCIENCES, INC.
STREET: 9410 KEY WEST AVENUE

```

; CITY: ROCKVILLE
; STATE: MD
; COUNTRY: US
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/832,488
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: BROOKES, ANDERS A.
; REGISTRATION NUMBER: 36,373
; REFERENCE/DOCKET NUMBER: PF266
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 857 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 82..756
; FEATURE:
; NAME/KEY: sig_peptide
; LOCATION: 82..130
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: 133..756
; US-08-832-488-1

```

```

Query Match      69.1%; Score 15.2; DB 4; Length 857;
Best Local Similarity 85.0%; Pred. No. 1.3e+02;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy  2  TGCAGGGGTTTGTCTGGGC 21
    |||||
Db  595 TGCAGCGGTTTGTCTGGAC 614

```

Search completed: April 18, 2003, 07:34:23
Job time : 32.5 secs

GenCore version 5.1.4 p5 4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 18, 2003, 06:35:31 ; Search time 91.3333 Seconds
(without alignments)
241.975 Million cell updates/sec

Title: US-09-270-437D-10
Perfect score: 22
Sequence: 1 ctgcaagggttttctgctggcg 22

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 639749 seqs, 502280378 residues
Total number of hits satisfying chosen parameters: 1279498

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA:*

- 1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
- 3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
- 4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
- 6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
- 8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
- 10: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
- 11: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
- 12: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
- 13: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
- 14: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	22	100.0	282	10	US-09-864-761-32671
C 2	22	100.0	588	10	US-09-864-761-16177
C 3	22	100.0	1740	10	US-09-735-705-347
C 4	22	100.0	1740	10	US-09-850-716A-347
C 5	22	100.0	1740	10	US-09-897-778-347
C 6	22	100.0	1743	10	US-09-897-778-447
C 7	22	100.0	1743	10	US-09-897-778-450
C 8	22	100.0	1764	10	US-09-850-716A-428
C 9	22	100.0	1764	10	US-09-897-778-428
C 10	22	100.0	4159	10	US-09-899-651-4
C 11	22	100.0	4181	10	US-09-735-705-175
C 12	22	100.0	4181	10	US-09-954-456-715
C 13	22	100.0	4181	10	US-09-850-716A-175
C 14	22	100.0	4181	10	US-09-897-778-175
C 15	17.2	78.2	302250	10	US-09-962-832-154
C 16	16.8	76.4	549	10	US-09-560-863-190
C 17	16.4	74.5	75899	10	US-09-854-883-243
C 18	16.2	73.6	296	10	US-09-294-093B-2223
C 19	16.2	73.6	368004	10	US-09-949-654-3

20	16.2	73.6	3309400	9	US-09-738-626-1	Sequence 1, Appli
21	15.8	71.8	261	10	US-09-878-574-15056	Sequence 15056, A
22	15.8	71.8	269	10	US-09-923-876-4150	Sequence 4150, Ap
23	15.8	71.8	272	10	US-09-878-574-14293	Sequence 14293, A
24	15.8	71.8	518	10	US-09-998-598-692	Sequence 692, App
25	15.8	71.8	2700	9	US-10-084-994-5	Sequence 5, Appli
C 26	15.8	71.8	3649	9	US-09-796-753-139	Sequence 139, App
C 27	15.6	70.9	377	10	US-09-867-701-2520	Sequence 2520, Ap
C 28	15.6	70.9	422	10	US-09-867-701-1470	Sequence 1470, Ap
C 29	15.6	70.9	426	9	US-09-738-626-1627	Sequence 1627, Ap
C 30	15.6	70.9	439	10	US-09-867-701-1485	Sequence 1485, Ap
C 31	15.6	70.9	454	10	US-09-867-701-5517	Sequence 5517, Ap
C 32	15.6	70.9	454	10	US-09-880-107-1265	Sequence 1265, Ap
C 33	15.6	70.9	454	10	US-09-967-768A-93	Sequence 93, Appli
C 34	15.6	70.9	747	9	US-09-738-626-2783	Sequence 2783, Ap
C 35	15.6	70.9	1455	9	US-09-738-626-1626	Sequence 1626, Ap
C 36	15.6	70.9	3487	9	US-10-114-170-164	Sequence 164, App
C 37	15.6	70.9	3489	10	US-09-815-242-7731	Sequence 7731, Ap
C 38	15.6	70.9	8522	10	US-09-817-181-3	Sequence 3, Appli
C 39	15.6	70.9	48908	9	US-10-114-170-137	Sequence 137, App
40	15.6	70.9	145831	10	US-09-969-708-79	Sequence 79, Appli
41	15.6	70.9	145831	10	US-09-954-456-2116	Sequence 2116, Ap
C 42	15.6	70.9	536165	9	US-09-939-964-1	Sequence 1, Appli
C 43	15.4	70.0	830	10	US-09-925-301-301	Sequence 301, App
C 44	15.2	69.1	163	10	US-09-833-381-687	Sequence 687, App
C 45	15.2	69.1	289	10	US-09-960-352-7109	Sequence 7109, Ap

ALIGNMENTS

RESULT 1

US-09-864-761-32671/c

Sequence 32671, Application US/09864761

Patent No. US20020048763A1

GENERAL INFORMATION:

APPLICANT: Penn, Sharron G.

APPLICANT: Rank, David R.

APPLICANT: Hanzel, David K.

APPLICANT: Chen, Wensheng

TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR

TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY

FILE REFERENCE: Aecomica-X-1

CURRENT APPLICATION NUMBER: US/09/864,761

CURRENT FILING DATE: 2001-05-23

PRIOR APPLICATION NUMBER: US 60/180,312

PRIOR FILING DATE: 2000-02-04

PRIOR APPLICATION NUMBER: US 60/207,456

PRIOR FILING DATE: 2000-05-26

PRIOR APPLICATION NUMBER: US 09/632,366

PRIOR FILING DATE: 2000-08-03

PRIOR APPLICATION NUMBER: GB 24263.6

PRIOR FILING DATE: 2000-10-04

PRIOR APPLICATION NUMBER: US 60/236,359

PRIOR FILING DATE: 2000-09-27

PRIOR APPLICATION NUMBER: PCT/US01/00666

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00667

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00664

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00669

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00665

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00668

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00663

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00662

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00661

PRIOR FILING DATE: 2001-01-30

;; PRIOR APPLICATION NUMBER: PCT/US01/00670
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: US 60/234,687
;; PRIOR FILING DATE: 2000-09-21
;; PRIOR APPLICATION NUMBER: US 09/608,408
;; PRIOR FILING DATE: 2000-06-30
;; PRIOR APPLICATION NUMBER: US 09/774,203
;; PRIOR FILING DATE: 2001-01-29
;; NUMBER OF SEQ ID NOS: 49117
;; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
;; SEQ ID NO 16177
;; LENGTH: 588
;; TYPE: DNA
;; ORGANISM: Homo sapiens
;; FEATURE:
;; OTHER INFORMATION: MAP TO AC021876.2
;; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 6.5
;; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 2.8
US-09-864-761-16177

Query Match 100.0%; Score 22; DB 10; Length 588;
Best Local Similarity 100.0%; Pred. No. 0.29; Mismatches 0; Indels 0; Gaps 0;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTGCAAGGGGTTTGTCTGGCG 22
|||
DB 117 CTGCAAGGGGTTTGTCTGGCG 96
|||

RESULT 3
US-09-735-705-347/c
;; Sequence 347, Application US/09735705
;; Patent No. US20020052329A1
;; GENERAL INFORMATION:
;; APPLICANT: Wang, Tongtong
;; APPLICANT: Fan, Liqun
;; APPLICANT: Kalos, Michael D.
;; APPLICANT: Bangur, Chaitanya S.
;; APPLICANT: Hosken, Nancy
;; APPLICANT: Fanger, Gary R.
;; APPLICANT: Li, Samuel X.
;; APPLICANT: Wang, Aijun
;; APPLICANT: Skeiky, Yasir A.W.
;; APPLICANT: Henderson, Robert A.
;; APPLICANT: McNeill, Patricia D.
;; APPLICANT: Fanger, Neil
;; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
;; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
;; FILE REFERENCE: 210121.455C14
;; CURRENT APPLICATION NUMBER: US/09/735,705
;; CURRENT FILING DATE: 2000-12-12
;; NUMBER OF SEQ ID NOS: 419
;; SOFTWARE: FastSEQ for Windows Version 3.0
;; SEQ ID NO 347
;; LENGTH: 1740
;; TYPE: DNA
;; ORGANISM: Homo sapiens
US-09-735-705-347

Query Match 100.0%; Score 22; DB 10; Length 1740;
Best Local Similarity 100.0%; Pred. No. 0.3; Mismatches 0; Indels 0; Gaps 0;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTGCAAGGGGTTTGTCTGGCG 22
|||
DB 498 CTGCAAGGGGTTTGTCTGGCG 477
|||

RESULT 4
US-09-850-716A-347/c
;; Sequence 347, Application US/09850716A
;; Patent No. US20020115139A1

;; PRIOR APPLICATION NUMBER: PCT/US01/00670
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: US 60/234,687
;; PRIOR FILING DATE: 2000-09-21
;; PRIOR APPLICATION NUMBER: US 09/608,408
;; PRIOR FILING DATE: 2000-06-30
;; PRIOR APPLICATION NUMBER: US 09/774,203
;; PRIOR FILING DATE: 2001-01-29
;; NUMBER OF SEQ ID NOS: 49117
;; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
;; SEQ ID NO 32671
;; LENGTH: 282
;; TYPE: DNA
;; ORGANISM: Homo sapiens
;; FEATURE:
;; OTHER INFORMATION: MAP TO AC021876.2
;; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 6.5
;; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 2.8
;; OTHER INFORMATION: SWISSPROT HIT: Q29125, EVALUE 3.20e-01
;; OTHER INFORMATION: NT HIT: U76705.1, EVALUE 0.00e+00
;; OTHER INFORMATION: EST_HUMAN HIT: BE545535.1, EVALUE 0.00e+00
US-09-864-761-32671

Query Match 100.0%; Score 22; DB 10; Length 282;
Best Local Similarity 100.0%; Pred. No. 0.28; Mismatches 0; Indels 0; Gaps 0;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTGCAAGGGGTTTGTCTGGCG 22
|||
DB 97 CTGCAAGGGGTTTGTCTGGCG 76
|||

RESULT 2
US-09-864-761-16177/c
;; Sequence 16177, Application US/09864761
;; Patent No. US20020048763A1
;; GENERAL INFORMATION:
;; APPLICANT: Penn, Sharon G.
;; APPLICANT: Rank, David R.
;; APPLICANT: Hanzel, David K.
;; APPLICANT: Chen, Wensheng
;; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
;; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
;; FILE REFERENCE: Aeomica-X-1
;; CURRENT APPLICATION NUMBER: US/09/864,761
;; CURRENT FILING DATE: 2001-05-23
;; PRIOR APPLICATION NUMBER: US 60/180,312
;; PRIOR FILING DATE: 2000-02-04
;; PRIOR APPLICATION NUMBER: US 60/207,456
;; PRIOR FILING DATE: 2000-05-26
;; PRIOR APPLICATION NUMBER: US 09/632,366
;; PRIOR FILING DATE: 2000-08-03
;; PRIOR APPLICATION NUMBER: GB 24263.6
;; PRIOR FILING DATE: 2000-10-04
;; PRIOR APPLICATION NUMBER: US 60/236,359
;; PRIOR FILING DATE: 2000-09-27
;; PRIOR APPLICATION NUMBER: PCT/US01/00666
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00667
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00664
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00669
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00665
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00668
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00663
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00662
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00661

GENERAL INFORMATION:
; APPLICANT: Kalos, Michael D.
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Retter, Marc W.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.455C15
; CURRENT APPLICATION NUMBER: US/09/850,716A
; CURRENT FILING DATE: 2001-05-07
; NUMBER OF SEQ ID NOS: 440
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 347:
; LENGTH: 1740
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-850-716A-347

Query Match 100.0%; Score 22; DB 10; Length 1740;
Best Local Similarity 100.0%; Pred. No. 0.3;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTGCAAGGGGTTTGTCTGGGCG 22
|||||
Db 498 CTGCAAGGGGTTTGTCTGGGCG 477

RESULT 5

US-09-897-778-347/c
; Sequence 347, Application US/09897778
; Patent No. US20020147143A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Marnerakis, Margarita
; APPLICANT: Fanger, Gary R.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Watanabe, Yoshihiro
; APPLICANT: Henderson, Robert A.
; APPLICANT: Peckham, David W.
; APPLICANT: Fanger, Neil
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.455C16
; CURRENT APPLICATION NUMBER: US/09/897,778
; CURRENT FILING DATE: 2001-06-28
; NUMBER OF SEQ ID NOS: 467
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 347
; LENGTH: 1740
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-897-778-347

Query Match 100.0%; Score 22; DB 10; Length 1740;
Best Local Similarity 100.0%; Pred. No. 0.3;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTGCAAGGGGTTTGTCTGGGCG 22
|||||
Db 498 CTGCAAGGGGTTTGTCTGGGCG 477

RESULT 6

US-09-897-778-447/c
; Sequence 447, Application US/09897778
; Patent No. US20020147143A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Marnerakis, Margarita
; APPLICANT: Fanger, Gary R.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Watanabe, Yoshihiro

; APPLICANT: Henderson, Robert A.
; APPLICANT: Peckham, David W.
; APPLICANT: Fanger, Neil
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.455C16
; CURRENT APPLICATION NUMBER: US/09/897,778
; CURRENT FILING DATE: 2001-06-28
; NUMBER OF SEQ ID NOS: 467
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 447
; LENGTH: 1743
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-897-778-447

Query Match 100.0%; Score 22; DB 10; Length 1743;
Best Local Similarity 100.0%; Pred. No. 0.3;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTGCAAGGGGTTTGTCTGGGCG 22
|||||
Db 498 CTGCAAGGGGTTTGTCTGGGCG 477

RESULT 7

US-09-897-778-450/c
; Sequence 450, Application US/09897778
; Patent No. US20020147143A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Marnerakis, Margarita
; APPLICANT: Fanger, Gary R.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Watanabe, Yoshihiro
; APPLICANT: Henderson, Robert A.
; APPLICANT: Peckham, David W.
; APPLICANT: Fanger, Neil
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.455C16
; CURRENT APPLICATION NUMBER: US/09/897,778
; CURRENT FILING DATE: 2001-06-28
; NUMBER OF SEQ ID NOS: 467
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 450
; LENGTH: 1743
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-897-778-450

Query Match 100.0%; Score 22; DB 10; Length 1743;
Best Local Similarity 100.0%; Pred. No. 0.3;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTGCAAGGGGTTTGTCTGGGCG 22
|||||
Db 498 CTGCAAGGGGTTTGTCTGGGCG 477

RESULT 8

US-09-850-716A-428/c
; Sequence 428, Application US/09850716A
; Patent No. US20020115139A1
; GENERAL INFORMATION:
; APPLICANT: Kalos, Michael D.
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Retter, Marc W.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.455C15
; CURRENT APPLICATION NUMBER: US/09/850,716A

us-09-270-437d-10.rnpb

Mon Apr 21 10:29:45 2003

```
; CURRENT FILING DATE: 2001-05-07
; NUMBER OF SEQ ID NOS: 440
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 428
; LENGTH: 1764
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-850-716A-428

Query Match      100.0%; Score 22; DB 10; Length 1764;
Best Local Similarity 100.0%; Pred. No. 0.3;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTGCAAGGGGTTTGTCTGGGCG 22
   |||||
DB 519 CTGCAAGGGGTTTGTCTGGGCG 498

RESULT 9
US-09-897-778-428/c
; Sequence 428, Application US/09897778
; Patent No. US20020147143A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Marnerakis, Margarita
; APPLICANT: Fanger, Gary R.
; APPLICANT: Vedwick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Watanabe, Yoshihiro
; APPLICANT: Henderson, Robert A.
; APPLICANT: Peckham, David W.
; APPLICANT: Fanger, Neil
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.455C16
; CURRENT APPLICATION NUMBER: US/09/897,778
; CURRENT FILING DATE: 2001-06-28
; NUMBER OF SEQ ID NOS: 467
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 428
; LENGTH: 1764
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-897-778-428

Query Match      100.0%; Score 22; DB 10; Length 1764;
Best Local Similarity 100.0%; Pred. No. 0.3;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTGCAAGGGGTTTGTCTGGGCG 22
   |||||
DB 519 CTGCAAGGGGTTTGTCTGGGCG 498

RESULT 10
US-09-899-651-4/c
; Sequence 4, Application US/09899651
; Patent No. US20020111470A1
; GENERAL INFORMATION:
; APPLICANT: Chen, Yao-Tsung
; APPLICANT: Gure, Ali
; APPLICANT: Tsang, Solam
; APPLICANT: Stockert, Elisabeth
; APPLICANT: Jager, Elke
; APPLICANT: Knuth, Alexander
; APPLICANT: Old, Lloyd J.
; TITLE OF INVENTION: Isolated Nucleic Acid Molecules Encoding Cancer
; TITLE OF INVENTION: Associated
; TITLE OF INVENTION: Antigen, The Antigens Per Se, And Uses Thereof
; FILE REFERENCE: LUD 5538
; CURRENT APPLICATION NUMBER: US/09/899,651
; CURRENT FILING DATE: 2001-07-06
; PRIOR APPLICATION NUMBER: US/09/061,709

; PRIOR FILING DATE: 1998-04-17
; NUMBER OF SEQ ID NOS: 8
; SEQ ID NO 4
; LENGTH: 4159
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
US-09-899-651-4

Query Match      100.0%; Score 22; DB 10; Length 4159;
Best Local Similarity 100.0%; Pred. No. 0.31;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTGCAAGGGGTTTGTCTGGGCG 22
   |||||
DB 748 CTGCAAGGGGTTTGTCTGGGCG 727

RESULT 11
US-09-735-705-175/c
; Sequence 175, Application US/09735705
; Patent No. US20020052329A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Fan, Liqun
; APPLICANT: Kalos, Michael D.
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Hosken, Nancy R.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Li, Samuel X.
; APPLICANT: Wang, Aljun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Henderson, Robert A.
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Fanger, Neil
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.455C14
; CURRENT APPLICATION NUMBER: US/09/735,705
; CURRENT FILING DATE: 2000-12-12
; NUMBER OF SEQ ID NOS: 419
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 175
; LENGTH: 4181
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (3347)
; OTHER INFORMATION: n=A,T,C or G
; NAME/KEY: unsure
; LOCATION: (3502)
; OTHER INFORMATION: n=A,T,C or G
; NAME/KEY: unsure
; LOCATION: (3506)
; OTHER INFORMATION: n=A,T,C or G
; NAME/KEY: unsure
; LOCATION: (3520)
; OTHER INFORMATION: n=A,T,C or G
; NAME/KEY: unsure
; LOCATION: (3538)
; OTHER INFORMATION: n=A,T,C or G
; NAME/KEY: unsure
; LOCATION: (3549)
; OTHER INFORMATION: n=A,T,C or G
; NAME/KEY: unsure
; LOCATION: (3646)
; OTHER INFORMATION: n=A,T,C or G
; NAME/KEY: unsure
; LOCATION: (3940)
; OTHER INFORMATION: n=A,T,C or G
; NAME/KEY: unsure
; LOCATION: (3968)
```

OTHER INFORMATION: n=A,T,C or G
 NAME/KEY: unsure
 LOCATION: (3974)
 OTHER INFORMATION: n=A,T,C or G
 NAME/KEY: unsure
 LOCATION: (4036)
 OTHER INFORMATION: n=A,T,C or G
 NAME/KEY: unsure
 LOCATION: (4056)
 OTHER INFORMATION: n=A,T,C or G
 NAME/KEY: unsure
 LOCATION: (4062)
 OTHER INFORMATION: n=A,T,C or G
 NAME/KEY: unsure
 LOCATION: (4080)
 OTHER INFORMATION: n=A,T,C or G
 NAME/KEY: unsure
 LOCATION: (4088)
 OTHER INFORMATION: n=A,T,C or G
 NAME/KEY: unsure
 LOCATION: (4115)
 OTHER INFORMATION: n=A,T,C or G
 US-09-735-705-175

Query Match 100.0%; Score 22; DB 10; Length 4181;
 Best Local Similarity 100.0%; Pred. No. 0.32;
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTGCAAGGGGTTTGTCTGGGCG 22
 Db 748 CTGCAAGGGGTTTGTCTGGGCG 727

RESULT 12

US-09-954-456-715/c
 Sequence 715, Application US/09954456
 Patent No. US20020115057A1
 GENERAL INFORMATION:
 APPLICANT: Young, Paul
 TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents Using Canc
 FILE OF INVENTION: Sets
 FILE REFERENCE: 689290-76
 CURRENT APPLICATION NUMBER: US/09/954,456
 PRIOR APPLICATION NUMBER: US/60/233,617
 PRIOR FILING DATE: 2001-09-18
 PRIOR APPLICATION NUMBER: US/60/234,052
 PRIOR FILING DATE: 2000-09-20
 PRIOR APPLICATION NUMBER: US/60/234,923
 PRIOR FILING DATE: 2000-09-25
 PRIOR APPLICATION NUMBER: US/60/235,134
 PRIOR FILING DATE: 2000-09-25
 PRIOR APPLICATION NUMBER: US/60/235,637
 PRIOR FILING DATE: 2000-09-26
 PRIOR APPLICATION NUMBER: US/60/235,638
 PRIOR FILING DATE: 2000-09-26
 PRIOR APPLICATION NUMBER: US/60/235,711
 PRIOR FILING DATE: 2000-09-27
 PRIOR APPLICATION NUMBER: US/60/235,720
 PRIOR FILING DATE: 2000-09-27
 PRIOR APPLICATION NUMBER: US/60/235,840
 PRIOR FILING DATE: 2000-09-27
 PRIOR APPLICATION NUMBER: US/60/235,863
 PRIOR FILING DATE: 2000-09-27
 NUMBER OF SEQ ID NOS: 2276
 SOFTWARE: Patent version 3.0
 SEQ ID NO 715
 LENGTH: 4181
 TYPE: DNA
 ORGANISM: Homo sapiens
 FEATURE:
 NAME/KEY: misc feature
 OTHER INFORMATION: n=a,t,g or c

US-09-954-456-715

Query Match 100.0%; Score 22; DB 10; Length 4181;
 Best Local Similarity 100.0%; Pred. No. 0.32;
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTGCAAGGGGTTTGTCTGGGCG 22
 Db 748 CTGCAAGGGGTTTGTCTGGGCG 727

RESULT 13

US-09-850-716A-175/c
 Sequence 175, Application US/09850716A
 Patent No. US20020115139A1
 GENERAL INFORMATION:
 APPLICANT: Kalos, Michael D.
 APPLICANT: McNeill, Patricia D.
 APPLICANT: Retter, Marc W.
 TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
 FILE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
 FILE REFERENCE: 210121.455C15
 CURRENT APPLICATION NUMBER: US/09/850,716A
 CURRENT FILING DATE: 2001-05-07
 NUMBER OF SEQ ID NOS: 440
 SOFTWARE: FastSeq for Windows Version 3.0
 SEQ ID NO 175
 LENGTH: 4181
 TYPE: DNA
 ORGANISM: Homo sapiens
 FEATURE:
 NAME/KEY: unsure
 LOCATION: (3347)
 OTHER INFORMATION: n=A,T,C or G
 NAME/KEY: unsure
 LOCATION: (3502)
 OTHER INFORMATION: n=A,T,C or G
 NAME/KEY: unsure
 LOCATION: (3538)
 OTHER INFORMATION: n=A,T,C or G
 NAME/KEY: unsure
 LOCATION: (3549)
 OTHER INFORMATION: n=A,T,C or G
 NAME/KEY: unsure
 LOCATION: (3520)
 OTHER INFORMATION: n=A,T,C or G
 NAME/KEY: unsure
 LOCATION: (3646)
 OTHER INFORMATION: n=A,T,C or G
 NAME/KEY: unsure
 LOCATION: (3940)
 OTHER INFORMATION: n=A,T,C or G
 NAME/KEY: unsure
 LOCATION: (3968)
 OTHER INFORMATION: n=A,T,C or G
 NAME/KEY: unsure
 LOCATION: (3974)
 OTHER INFORMATION: n=A,T,C or G
 NAME/KEY: unsure
 LOCATION: (4036)
 OTHER INFORMATION: n=A,T,C or G
 NAME/KEY: unsure
 LOCATION: (4056)
 OTHER INFORMATION: n=A,T,C or G
 NAME/KEY: unsure
 LOCATION: (4062)
 OTHER INFORMATION: n=A,T,C or G
 NAME/KEY: unsure
 LOCATION: (4080)
 OTHER INFORMATION: n=A,T,C or G

us-09-270-437d-10.rnpb

Mon Apr 21 10:29:45 2003

```

; NAME/KEY: unsure
; LOCATION: (4088)
; OTHER INFORMATION: n=A,T,C or G
; NAME/KEY: unsure
; LOCATION: (4115)
; OTHER INFORMATION: n=A,T,C or G
US-09-850-716A-175

Query Match      100.0%; Score 22; DB 10; Length 4181;
Best Local Similarity 100.0%; Pred. No. 0.32;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTGCAAGGGGTTTGTCTGGCG 22
    |||||
Db 748 CTGCAAGGGGTTTGTCTGGCG 727

; NUMBER OF SEQ ID NOS: 259
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 154
; LENGTH: 302250
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-962-832-154

Query Match      78.2%; Score 17.2; DB 10; Length 302250;
Best Local Similarity 86.4%; Pred. No. 65;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 CTGCAAGGGGTTTGTCTGGCG 22
    |||||
Db 91459 CTGCAAGGGGTACAGCTGGCG 91438

Search completed: April 18, 2003, 10:14:02
Job time : 149.5 secs

```

```

; NAME/KEY: unsure
; LOCATION: (4088)
; OTHER INFORMATION: n=A,T,C or G
; NAME/KEY: unsure
; LOCATION: (4115)
; OTHER INFORMATION: n=A,T,C or G
US-09-850-716A-175

Query Match      100.0%; Score 22; DB 10; Length 4181;
Best Local Similarity 100.0%; Pred. No. 0.32;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTGCAAGGGGTTTGTCTGGCG 22
    |||||
Db 748 CTGCAAGGGGTTTGTCTGGCG 727

RESULT 14
US-09-897-778-175/c
; Sequence 175, Application US/09897778
; Patent No. US20020147143A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Marnerakis, Margarita
; APPLICANT: Fanger, Gary R.
; APPLICANT: Vedwick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Watanabe, Yoshihiro
; APPLICANT: Henderson, Robert A.
; APPLICANT: Peckham, David W.
; APPLICANT: Fanger, Neil
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.455C16
; CURRENT APPLICATION NUMBER: US/09/897,778
; CURRENT FILING DATE: 2001-06-28
; NUMBER OF SEQ ID NOS: 467
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 175
; LENGTH: 4181
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 3347, 3502, 3506, 3520, 3538, 3549, 3646, 3940, 3968, 3974,
; LOCATION: 4036, 4056, 4062, 4080, 4088, 4115
; OTHER INFORMATION: n = A,T,C or G
US-09-897-778-175

```

```

Query Match      100.0%; Score 22; DB 10; Length 4181;
Best Local Similarity 100.0%; Pred. No. 0.32;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 CTGCAAGGGGTTTGTCTGGCG 22
    |||||
Db 748 CTGCAAGGGGTTTGTCTGGCG 727

```

```

RESULT 15
US-09-962-832-154/c
; Sequence 154, Application US/09962832
; Patent No. US20020110821A1
; GENERAL INFORMATION:
; APPLICANT: Ebner, Reinhard
; TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using Signatu
; TITLE OF INVENTION: Sets
; FILE REFERENCE: 689290-74
; CURRENT APPLICATION NUMBER: US/09/962,832
; CURRENT FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US/60/235,077
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US/60/235,280
; PRIOR FILING DATE: 2000-09-25

```

GenCore version 5.1.4 p5 4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 18, 2003, 05:48:17 ; Search time 1211 Seconds
(without alignments)
456.759 Million cell updates/sec

Title: US-09-270-437D-10

Perfect score: 22

Sequence: 1 ctgcaagggttttctgctggcg 22

Scoring table: IDENTITY NUC

Gapop 10_0 , Gapext 1.0

Searched: 24791104 seqs, 12571243825 residues

Total number of hits satisfying chosen parameters: 49582208

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Pending Patents_NA_Main:

- 1: /cgn2_6/ptodata/1/pna/pctus COMB.seq.*
- 2: /cgn2_6/ptodata/1/pna/US06 COMB.seq.*
- 3: /cgn2_6/ptodata/1/pna/US07 COMB.seq.*
- 4: /cgn2_6/ptodata/1/pna/US08 COMB.seq.*
- 5: /cgn2_6/ptodata/1/pna/US081 COMB.seq.*
- 6: /cgn2_6/ptodata/1/pna/US082 COMB.seq.*
- 7: /cgn2_6/ptodata/1/pna/US083 COMB.seq.*
- 8: /cgn2_6/ptodata/1/pna/US084 COMB.seq.*
- 9: /cgn2_6/ptodata/1/pna/US085 COMB.seq.*
- 10: /cgn2_6/ptodata/1/pna/US086 COMB.seq.*
- 11: /cgn2_6/ptodata/1/pna/US087 COMB.seq.*
- 12: /cgn2_6/ptodata/1/pna/US088 COMB.seq.*
- 13: /cgn2_6/ptodata/1/pna/US089 COMB.seq.*
- 14: /cgn2_6/ptodata/1/pna/US090 COMB.seq.*
- 15: /cgn2_6/ptodata/1/pna/US091 COMB.seq.*
- 16: /cgn2_6/ptodata/1/pna/US092 COMB.seq.*
- 17: /cgn2_6/ptodata/1/pna/US093 COMB.seq.*
- 18: /cgn2_6/ptodata/1/pna/US094 COMB.seq.*
- 19: /cgn2_6/ptodata/1/pna/US095A COMB.seq.*
- 20: /cgn2_6/ptodata/1/pna/US095B COMB.seq.*
- 21: /cgn2_6/ptodata/1/pna/US095C COMB.seq.*
- 22: /cgn2_6/ptodata/1/pna/US095D COMB.seq.*
- 23: /cgn2_6/ptodata/1/pna/US096A COMB.seq.*
- 24: /cgn2_6/ptodata/1/pna/US096B COMB.seq.*
- 25: /cgn2_6/ptodata/1/pna/US096C COMB.seq.*
- 26: /cgn2_6/ptodata/1/pna/US096D COMB.seq.*
- 27: /cgn2_6/ptodata/1/pna/US096E COMB.seq.*
- 28: /cgn2_6/ptodata/1/pna/US097A COMB.seq.*
- 29: /cgn2_6/ptodata/1/pna/US097B COMB.seq.*
- 30: /cgn2_6/ptodata/1/pna/US097C COMB.seq.*
- 31: /cgn2_6/ptodata/1/pna/US098A COMB.seq.*
- 32: /cgn2_6/ptodata/1/pna/US098B COMB.seq.*
- 33: /cgn2_6/ptodata/1/pna/US098C COMB.seq.*
- 34: /cgn2_6/ptodata/1/pna/US099A COMB.seq.*
- 35: /cgn2_6/ptodata/1/pna/US099B COMB.seq.*
- 36: /cgn2_6/ptodata/1/pna/US099C COMB.seq.*
- 37: /cgn2_6/ptodata/1/pna/US099D COMB.seq.*
- 38: /cgn2_6/ptodata/1/pna/US100A COMB.seq.*
- 39: /cgn2_6/ptodata/1/pna/US100B COMB.seq.*
- 40: /cgn2_6/ptodata/1/pna/US101A COMB.seq.*
- 41: /cgn2_6/ptodata/1/pna/US101B COMB.seq.*
- 42: /cgn2_6/ptodata/1/pna/US102A COMB.seq.*
- 43: /cgn2_6/ptodata/1/pna/US102B COMB.seq.*

- 44: /cgn2_6/ptodata/1/pna/US6000 COMB.seq.*
- 45: /cgn2_6/ptodata/1/pna/US6001 COMB.seq.*
- 46: /cgn2_6/ptodata/1/pna/US6002 COMB.seq.*
- 47: /cgn2_6/ptodata/1/pna/US6003 COMB.seq.*
- 48: /cgn2_6/ptodata/1/pna/US6004 COMB.seq.*
- 49: /cgn2_6/ptodata/1/pna/US6005 COMB.seq.*
- 50: /cgn2_6/ptodata/1/pna/US6006 COMB.seq.*
- 51: /cgn2_6/ptodata/1/pna/US6007 COMB.seq.*
- 52: /cgn2_6/ptodata/1/pna/US6008 COMB.seq.*
- 53: /cgn2_6/ptodata/1/pna/US6009 COMB.seq.*
- 54: /cgn2_6/ptodata/1/pna/US6010 COMB.seq.*
- 55: /cgn2_6/ptodata/1/pna/US6011 COMB.seq.*
- 56: /cgn2_6/ptodata/1/pna/US6012 COMB.seq.*
- 57: /cgn2_6/ptodata/1/pna/US6013 COMB.seq.*
- 58: /cgn2_6/ptodata/1/pna/US6014 COMB.seq.*
- 59: /cgn2_6/ptodata/1/pna/US6015 COMB.seq.*
- 60: /cgn2_6/ptodata/1/pna/US6016 COMB.seq.*
- 61: /cgn2_6/ptodata/1/pna/US6017 COMB.seq.*
- 62: /cgn2_6/ptodata/1/pna/US6018 COMB.seq.*
- 63: /cgn2_6/ptodata/1/pna/US6019 COMB.seq.*
- 64: /cgn2_6/ptodata/1/pna/US6020 COMB.seq.*
- 65: /cgn2_6/ptodata/1/pna/US6021 COMB.seq.*
- 66: /cgn2_6/ptodata/1/pna/US6022 COMB.seq.*
- 67: /cgn2_6/ptodata/1/pna/US6023 COMB.seq.*
- 68: /cgn2_6/ptodata/1/pna/US6024 COMB.seq.*
- 69: /cgn2_6/ptodata/1/pna/US6025 COMB.seq.*
- 70: /cgn2_6/ptodata/1/pna/US6026 COMB.seq.*
- 71: /cgn2_6/ptodata/1/pna/US6027 COMB.seq.*
- 72: /cgn2_6/ptodata/1/pna/US6028 COMB.seq.*
- 73: /cgn2_6/ptodata/1/pna/US6029 COMB.seq.*
- 74: /cgn2_6/ptodata/1/pna/US6030 COMB.seq.*
- 75: /cgn2_6/ptodata/1/pna/US6031 COMB.seq.*
- 76: /cgn2_6/ptodata/1/pna/US6032 COMB.seq.*
- 77: /cgn2_6/ptodata/1/pna/US6033 COMB.seq.*
- 78: /cgn2_6/ptodata/1/pna/US6034 COMB.seq.*
- 79: /cgn2_6/ptodata/1/pna/US6035 COMB.seq.*
- 80: /cgn2_6/ptodata/1/pna/US6036 COMB.seq.*
- 81: /cgn2_6/ptodata/1/pna/US6037 COMB.seq.*
- 82: /cgn2_6/ptodata/1/pna/US6038 COMB.seq.*
- 83: /cgn2_6/ptodata/1/pna/US6039 COMB.seq.*
- 84: /cgn2_6/ptodata/1/pna/US6040 COMB.seq.*
- 85: /cgn2_6/ptodata/1/pna/US6041 COMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	22	100.0	282	1	PCT-US01-00663-25064 Sequence 25064, A
C 2	22	100.0	282	33	US-09-864-761-32671 Sequence 32671, A
C 3	22	100.0	282	42	US-10-203-134-24949 Sequence 24949, A
C 4	22	100.0	282	42	US-10-203-137-25064 Sequence 25064, A
C 5	22	100.0	285	29	US-09-758-442-323 Sequence 323, App
C 6	22	100.0	285	42	US-10-217-623-323 Sequence 323, App
C 7	22	100.0	301	14	US-09-076-898-395 Sequence 395, App
C 8	22	100.0	301	14	US-09-076-898A-395 Sequence 395, App
C 9	22	100.0	301	34	US-09-912-292-11611 Sequence 11611, A
C 10	22	100.0	369	20	US-08-668-236-703 Sequence 703, App
C 11	22	100.0	369	20	US-09-534-846B-34372 Sequence 34372, A
C 12	22	100.0	404	14	US-09-076-897-1393 Sequence 1393, App
C 13	22	100.0	404	14	US-09-076-897A-1393 Sequence 1393, App
C 14	22	100.0	404	14	US-09-076-897B-1393 Sequence 1393, App
C 15	22	100.0	404	14	US-09-076-897C-1393 Sequence 1393, App
C 16	22	100.0	588	1	PCT-US01-00663-12084 Sequence 12084, A
C 17	22	100.0	588	33	US-09-864-761-16177 Sequence 16177, A
C 18	22	100.0	588	42	US-10-203-134-12077 Sequence 12077, A
C 19	22	100.0	588	42	US-10-203-137-12084 Sequence 12084, A
C 20	22	100.0	843	17	US-09-399-932-4493 Sequence 4493, App
C 21	22	100.0	1740	1	PCT-US01-47576-347 Sequence 347, App

Page 2

```

Matches      22;  Conservative      0;  Mismatches      0;  Indels      0;  Gaps      0;

OY      1      CTGCAAGGGGTTTTGCTGGGCG      22
          |||||
Db      97      CTGCAAGGGGTTTTGCTGGGCG      76

RESULT 2
US-09-864-761-32671/c
; Sequence 32671, Application US/09864761
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
; FILE REFERENCE: Aeomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6

```

; PRIOR APPLICATION NUMBER: GB 24263:6

```

, PRIOR FILING DATE: 2000-09-27
, PRIOR APPLICATION NUMBER: PCT/US01/006666
, PRIOR FILING DATE: 2001-01-30
, PRIOR APPLICATION NUMBER: PCT/US01/006667
, PRIOR FILING DATE: 2001-01-30
, PRIOR APPLICATION NUMBER: PCT/US01/006664
, PRIOR FILING DATE: 2001-01-30
, PRIOR APPLICATION NUMBER: PCT/US01/006669
, PRIOR FILING DATE: 2001-01-30
, PRIOR APPLICATION NUMBER: PCT/US01/006665
, PRIOR FILING DATE: 2001-01-30
, PRIOR APPLICATION NUMBER: PCT/US01/006668
, PRIOR FILING DATE: 2001-01-30
, PRIOR APPLICATION NUMBER: PCT/US01/006663
, PRIOR FILING DATE: 2001-01-30
, PRIOR APPLICATION NUMBER: PCT/US01/006662
, PRIOR FILING DATE: 2001-01-30
, PRIOR APPLICATION NUMBER: PCT/US01/006661
, PRIOR FILING DATE: 2001-01-30
, PRIOR APPLICATION NUMBER: PCT/US01/006670
, PRIOR FILING DATE: 2001-01-30
, PRIOR APPLICATION NUMBER: US 60/234,687
, PRIOR FILING DATE: 2000-09-21
, PRIOR APPLICATION NUMBER: US 09/608,408
, PRIOR FILING DATE: 2000-06-30
, PRIOR APPLICATION NUMBER: US 09/774,203
, PRIOR FILING DATE: 2001-01-29
, NUMBER OF SEQ ID NOS: 49117
, SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
, SEQ ID NO 32671
, LENGTH: 282
, TYPE: DNA
, ORGANISM: Homo sapiens
, FEATURE:
, OTHER INFORMATION: MAP TO AC021876.2
, OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 6.5
, OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 2.8
, OTHER INFORMATION: SWISSPROT HIT: Q29125, EVALUE 3.20e-01
, OTHER INFORMATION: NT HIT: U76705.1, EVALUE 0.00e+00
, OTHER INFORMATION: EST_HUMAN HIT: BE545535.1, EVALUE 0.00e+00
, US-09-864-761-32671

```

Query Match	100.0%;	Score 22;	DB 33;	Length 282;
Best Local Similarity	100.0%;	Pred. No. 5.4;		

```
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 CTGCAAGGGGTTTGTGGGCG 22
Db 97 CTGCAAGGGGTTTGTGGGCG 76

RESULT 3
US-10-203-134-24949/c
; Sequence 24949, Application US/10203134
; GENERAL INFORMATION:
; APPLICANT: Molecular Dynamics, Inc.
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; TITLE OF INVENTION: ANALYSIS OF GENE EXPRESSION IN HUMAN BONE MARROW
; FILE REFERENCE: PB 0004 WO 6
; CURRENT APPLICATION NUMBER: US/10/203,134
; PRIOR FILING DATE: 2002-08-02
; PRIOR FILING DATE: 04 February 2000 (04.02.00)
; PRIOR FILING DATE: 04 February 2000 (04.02.00)
; PRIOR FILING DATE: 26 May 2000 (26.05.00)
; PRIOR FILING DATE: 03 August 2000 (03.08.00)
; PRIOR FILING DATE: 03 August 2000 (03.08.00)
; PRIOR FILING DATE: 03 October 2000 (03.10.00)
; PRIOR FILING DATE: 27 September 2000 (27.09.00)
; PRIOR FILING DATE: 21 September 2000 (21.09.00)
; PRIOR FILING DATE: 21 September 2000 (21.09.00)
; PRIOR FILING DATE: 30 June 2000 (30.06.00)
; NUMBER OF SEQ ID NOS: 38628
; SOFTWARE: Molecular Dynamics Sequence Listing Engine
; SEQ ID NO 24949
; LENGTH: 282
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC021876.2
; FEATURE:
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 6.5
; FEATURE:
; OTHER INFORMATION: SWISSPROT HIT: Q29125, EVALUE 3.20e-01
; FEATURE:
; OTHER INFORMATION: NT HIT: U76705.1, EVALUE 0.00e+00
; OTHER INFORMATION: EST_HUMAN HIT: BE545535.1, EVALUE 0.00e+00
US-10-203-134-24949
Query Match 100.0%; Score 22; DB 42; Length 282;
Best Local Similarity 100.0%; Pred. No. 5.4;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 CTGCAAGGGGTTTGTGGGCG 22
Db 97 CTGCAAGGGGTTTGTGGGCG 76

RESULT 4
US-10-203-137-25064/c
; Sequence 25064, Application US/10203137
; GENERAL INFORMATION:
; APPLICANT: Molecular Dynamics, Inc.
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; TITLE OF INVENTION: ANALYSIS OF GENE EXPRESSION IN HUMAN PLACENTA
```

```
FILE REFERENCE: PB 0004 WO 7
; CURRENT APPLICATION NUMBER: US/10/203,137
; CURRENT FILING DATE: 2002-08-02
; PRIOR FILING DATE: 04 February 2000 (04.02.00)
; PRIOR FILING DATE: 04 February 2000 (04.02.00)
; PRIOR FILING DATE: 26 May 2000 (26.05.00)
; PRIOR FILING DATE: 03 August 2000 (03.08.00)
; PRIOR FILING DATE: 03 August 2000 (03.08.00)
; PRIOR FILING DATE: 03 October 2000 (03.10.00)
; PRIOR FILING DATE: 27 September 2000 (27.09.00)
; PRIOR FILING DATE: 21 September 2000 (21.09.00)
; PRIOR FILING DATE: 21 September 2000 (21.09.00)
; PRIOR FILING DATE: 30 June 2000 (30.06.00)
; NUMBER OF SEQ ID NOS: 38837
; SOFTWARE: Molecular Dynamics Sequence Listing Engine
; SEQ ID NO 25064
; LENGTH: 282
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC021876.2
; FEATURE:
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 2.8
; FEATURE:
; OTHER INFORMATION: SWISSPROT HIT: Q29125, EVALUE 3.20e-01
; FEATURE:
; OTHER INFORMATION: NT HIT: U76705.1, EVALUE 0.00e+00
; OTHER INFORMATION: EST_HUMAN HIT: BE545535.1, EVALUE 0.00e+00
US-10-203-137-25064
Query Match 100.0%; Score 22; DB 42; Length 282;
Best Local Similarity 100.0%; Pred. No. 5.4;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 CTGCAAGGGGTTTGTGGGCG 22
Db 97 CTGCAAGGGGTTTGTGGGCG 76

RESULT 5
US-09-758-442-323/c
; Sequence 323, Application US/09758442
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PM031
; CURRENT APPLICATION NUMBER: US/09/758,442
; CURRENT FILING DATE: 2001-01-11
; PRIOR FILING DATE: 2000-01-31
; PRIOR FILING DATE: 2000-01-31
; PRIOR FILING DATE: 2000-01-31
; PRIOR FILING DATE: 2000-02-04
; NUMBER OF SEQ ID NOS: 710
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 323
; LENGTH: 285
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (109)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (112)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-758-442-323
Query Match 100.0%; Score 22; DB 29; Length 285;
```

Best Local Similarity 100.0%; Pred. No. 5.4; Indels 0; Gaps 0; Mismatches 0; Matches 22; Conservative 0;

Qy 1 CTGCAAGGGGTTTCTGGGCG 22
Db 155 CTGCAAGGGGTTTCTGGGCG 134

RESULT 6
US-10-217-623-323/c
Sequence 323, Application US/10217623
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
FILE REFERENCE: PM031CIN
CURRENT APPLICATION NUMBER: US/10/217,623
PRIORITY FILING DATE: 2002-08-14
PRIOR APPLICATION NUMBER: 09/758,442
PRIOR FILING DATE: 2001-01-11
PRIOR APPLICATION NUMBER: 60/179,065
PRIOR FILING DATE: 2000-01-31
PRIOR APPLICATION NUMBER: 60/180,628
PRIOR FILING DATE: 2000-02-04
NUMBER OF SEQ ID NOS: 710
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 323
TYPE: DNA
LENGTH: 285
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
LOCATION: (109)
OTHER INFORMATION: n equals a,t,g, or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (112)
OTHER INFORMATION: n equals a,t,g, or c
US-10-217-623-323

Query Match 100.0%; Score 22; DB 42; Length 285;
Best Local Similarity 100.0%; Pred. No. 5.4;
Matches 22; Conservative 0; Indels 0; Gaps 0; Mismatches 0;

Qy 1 CTGCAAGGGGTTTCTGGGCG 22
Db 155 CTGCAAGGGGTTTCTGGGCG 134

RESULT 7
US-09-076-898-395/c
Sequence 395, Application US/09076898
GENERAL INFORMATION:
APPLICANT: Craig A. Rosen, et. al.
TITLE OF INVENTION: Human Genes, Sequences, and Expression Products - P031
NUMBER OF SEQUENCES: 2204
CORRESPONDENCE ADDRESS:
ADDRESSEE: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/076,898
FILING DATE: May 13, 1998
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/047,004

FILING DATE: May 13, 1997
ATTORNEY/AGENT INFORMATION:
NAME: Kenley K. Hoover
REGISTRATION NUMBER: 40,302
REFERENCE/DOCKET NUMBER: P031
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 309-8504
TELEFAX: (301) 309-8439
INFORMATION FOR SEQ ID NO: 395:
SEQUENCE CHARACTERISTICS:
LENGTH: 301 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
US-09-076-898-395

Query Match 100.0%; Score 22; DB 14; Length 301;
Best Local Similarity 100.0%; Pred. No. 5.5;
Matches 22; Conservative 0; Indels 0; Gaps 0; Mismatches 0;

Qy 1 CTGCAAGGGGTTTCTGGGCG 22
Db 162 CTGCAAGGGGTTTCTGGGCG 141

RESULT 8
US-09-076-898A-395/c
Sequence 395, Application US/09076898A
GENERAL INFORMATION:
APPLICANT: Rosen, Craig A., et. al.
TITLE OF INVENTION: Human Genes, Sequences, and Expression Products 31
FILE REFERENCE: PO-31
CURRENT APPLICATION NUMBER: US/09/076,898A
CURRENT FILING DATE: 1998-05-13
PRIOR APPLICATION NUMBER: 60/047,004
PRIOR FILING DATE: 1997-05-13
NUMBER OF SEQ ID NOS: 2204
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 395
TYPE: DNA
LENGTH: 301
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
LOCATION: (6)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc_feature
LOCATION: (115)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc_feature
LOCATION: (118)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc_feature
LOCATION: (261)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc_feature
LOCATION: (263)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc_feature
LOCATION: (264)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc_feature
LOCATION: (284)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc_feature
LOCATION: (289)
OTHER INFORMATION: n equals a,t,g, or c
US-09-076-898A-395

Query Match 100.0%; Score 22; DB 14; Length 301;
Best Local Similarity 100.0%; Pred. No. 5.5;
Matches 22; Conservative 0; Indels 0; Gaps 0; Mismatches 0;

Qy 1 CTGCAAGGGGTTTGTCTGGCG 22
Db 162 CTGCAAGGGGTTTGTCTGGCG 141

RESULT 9
US-09-912-292-11611/c
; Sequence 11611, Application US/09912292
; GENERAL INFORMATION:
; APPLICANT: Rosen, et. al.
; TITLE OF INVENTION: Human Genes, Sequences, and Expression Products 101
; FILE REFERENCE: PO-101
; CURRENT APPLICATION NUMBER: US/09/912,292
; CURRENT FILING DATE: 2001-07-26
; PRIOR APPLICATION NUMBER: 08/103,746
; PRIOR FILING DATE: 1993-08-09
; PRIOR APPLICATION NUMBER: 09/859,417
; PRIOR FILING DATE: 2001-05-18
; PRIOR APPLICATION NUMBER: 08/103,743
; PRIOR FILING DATE: 1993-08-09
; PRIOR APPLICATION NUMBER: 09/741,827
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 09/813,154
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 08/196,482
; PRIOR FILING DATE: 1994-02-15
; PRIOR APPLICATION NUMBER: 09/783,587
; PRIOR FILING DATE: 2001-02-15
; PRIOR APPLICATION NUMBER: 08/196,481
; PRIOR FILING DATE: 1994-02-15
; PRIOR APPLICATION NUMBER: 08/220,662
; PRIOR FILING DATE: 1994-03-31
; PRIOR APPLICATION NUMBER: 08/220,661
; PRIOR FILING DATE: 1994-03-31
; PRIOR APPLICATION NUMBER: 08/275,627
; PRIOR FILING DATE: 1994-07-15
; PRIOR APPLICATION NUMBER: 08/345,704
; PRIOR FILING DATE: 1994-11-21
; PRIOR APPLICATION NUMBER: 09/789,662
; PRIOR FILING DATE: 2001-05-18
; PRIOR APPLICATION NUMBER: 08/401,881
; PRIOR FILING DATE: 1995-03-10
; PRIOR APPLICATION NUMBER: 09/782,161
; PRIOR FILING DATE: 2001-02-14
; PRIOR APPLICATION NUMBER: 08/790,776
; PRIOR FILING DATE: 1997-01-30
; PRIOR APPLICATION NUMBER: 60/010,803
; PRIOR FILING DATE: 1996-01-30
; PRIOR APPLICATION NUMBER: 08/799,180
; PRIOR FILING DATE: 1997-02-12
; PRIOR APPLICATION NUMBER: 60/011,618
; PRIOR FILING DATE: 1996-02-13
; PRIOR APPLICATION NUMBER: 08/803,609
; PRIOR FILING DATE: 1997-02-21
; PRIOR APPLICATION NUMBER: 60/011,985
; PRIOR FILING DATE: 1996-02-21
; PRIOR APPLICATION NUMBER: 08/971,050
; PRIOR FILING DATE: 1997-11-14
; PRIOR APPLICATION NUMBER: 60/031,038
; PRIOR FILING DATE: 1996-11-18
; PRIOR APPLICATION NUMBER: 08/975,985
; PRIOR FILING DATE: 1997-11-21
; PRIOR APPLICATION NUMBER: 60/031,221
; PRIOR FILING DATE: 1996-11-25
; PRIOR APPLICATION NUMBER: 08/985,366
; PRIOR FILING DATE: 1997-12-04
; PRIOR APPLICATION NUMBER: 09/842,827
; PRIOR FILING DATE: 2001-04-27
; PRIOR APPLICATION NUMBER: 60/032,782
; PRIOR FILING DATE: 1996-12-06
; PRIOR APPLICATION NUMBER: 09/076,898
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/047,004

; PRIOR FILING DATE: 1997-05-13
; PRIOR APPLICATION NUMBER: 09/078,614
; PRIOR FILING DATE: 1998-05-14
; PRIOR APPLICATION NUMBER: 60/046,498
; PRIOR FILING DATE: 1997-05-14
; PRIOR APPLICATION NUMBER: 09/189,834
; PRIOR FILING DATE: 1998-11-12
; PRIOR APPLICATION NUMBER: 60/065,454
; PRIOR FILING DATE: 1997-11-12
; PRIOR APPLICATION NUMBER: 09/366,690
; PRIOR FILING DATE: 1999-08-04
; PRIOR APPLICATION NUMBER: 60/095,484
; PRIOR FILING DATE: 1998-08-05
; PRIOR APPLICATION NUMBER: 09/342,216
; PRIOR FILING DATE: 1999-06-29
; PRIOR APPLICATION NUMBER: 60/091,171
; PRIOR FILING DATE: 1998-06-30
; PRIOR APPLICATION NUMBER: 09/342,061
; PRIOR FILING DATE: 1999-06-29
; PRIOR APPLICATION NUMBER: 60/091,170
; PRIOR FILING DATE: 1998-06-30
; PRIOR APPLICATION NUMBER: 09/487,572
; PRIOR FILING DATE: 2000-01-19
; PRIOR APPLICATION NUMBER: 60/116,678
; PRIOR FILING DATE: 1999-01-21
; PRIOR APPLICATION NUMBER: 09/497,772
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: 60/118,784
; PRIOR FILING DATE: 1999-02-05
; PRIOR APPLICATION NUMBER: 09/504,577
; PRIOR FILING DATE: 2000-02-15
; PRIOR APPLICATION NUMBER: 60/120,434
; PRIOR FILING DATE: 1999-02-16
; PRIOR APPLICATION NUMBER: 09/517,011
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: 60/122,447
; PRIOR FILING DATE: 1999-03-02
; PRIOR APPLICATION NUMBER: 09/628,857
; PRIOR FILING DATE: 2000-07-27
; PRIOR APPLICATION NUMBER: 60/145,956
; PRIOR FILING DATE: 1999-07-28
; PRIOR APPLICATION NUMBER: 09/628,858
; PRIOR FILING DATE: 2000-07-27
; PRIOR APPLICATION NUMBER: 60/145,955
; PRIOR FILING DATE: 1999-07-28
; PRIOR APPLICATION NUMBER: 09/741,033
; PRIOR FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 09/773,518
; PRIOR FILING DATE: 2001-02-02
; PRIOR APPLICATION NUMBER: 60/171,621
; PRIOR FILING DATE: 1999-12-23
; PRIOR APPLICATION NUMBER: 09/796,765
; PRIOR FILING DATE: 2001-03-02
; PRIOR APPLICATION NUMBER: 09/867,682
; PRIOR FILING DATE: 2001-05-31
; PRIOR APPLICATION NUMBER: 60/187,016
; PRIOR FILING DATE: 2000-03-06
; PRIOR APPLICATION NUMBER: 60/225,269
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/225,629
; PRIOR FILING DATE: 2000-08-15
; NUMBER OF SEQ ID NOS: 55553
; SEQ ID NO 11611
; LENGTH: 301
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (6)..(6)
; OTHER INFORMATION: n is equal to a,t,g, or c
; NAME/KEY: misc.feature
; LOCATION: (115)..(115)
; OTHER INFORMATION: n is equal to a,t,g, or c

NAME/KEY: misc feature
LOCATION: (118)..(118)
OTHER INFORMATION: n is equal to a,t,g, or c
NAME/KEY: misc feature
LOCATION: (261)..(261)
OTHER INFORMATION: n is equal to a,t,g, or c
NAME/KEY: misc feature
LOCATION: (263)..(264)
OTHER INFORMATION: n is equal to a,t,g, or c
NAME/KEY: misc feature
LOCATION: (284)..(284)
OTHER INFORMATION: n is equal to a,t,g, or c
NAME/KEY: misc feature
LOCATION: (289)..(289)
OTHER INFORMATION: n is equal to a,t,g, or c
US-09-912-232-11611

Query Match 100.0%; Score 22; DB 34; Length 301;
Best Local Similarity 100.0%; Pred. No. 5.5; 0; Indels 0; Gaps 0;
Matches 22; Conservative 0; Mismatches 0

QY 1 CTGCAAGGGGTTTGTCTGGCG 22
DB 162 CTGCAAGGGGTTTGTCTGGCG 141

RESULT 10
US-08-668-236-703/c
Sequence 703, Application US/08668236
GENERAL INFORMATION:
APPLICANT: Stuart, Susan G.
APPLICANT: Delegeane, Angelo M.
APPLICANT: Bills, Pamela Kay
APPLICANT: Pham, Mino T.
APPLICANT: Altus, Christina T.
APPLICANT: Hillman, Jennifer L.
APPLICANT: Akersblom, Ingrid E.
TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM A
TITLE OF INVENTION: HUMAN NEURONAL PRECURSOR CELL LINE
NUMBER OF SEQUENCES: 1403
CORRESPONDENCE ADDRESS:
ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
STREET: 3330 HILLVIEW AVENUE
CITY: PALO ALTO
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/668,236
FILING DATE:
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: LUTHER, BARBARA J.
REGISTRATION NUMBER: 33954
REFERENCE/DOCKET NUMBER: PD-0038P
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 855-0555
TELEFAX: (415) 852-0195
INFORMATION FOR SEQ ID NO: 703:
SEQUENCE CHARACTERISTICS:
LENGTH: 369 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
IMMEDIATE SOURCE:
CLONE: 268739
US-08-668-236-703

Query Match 100.0%; Score 22; DB 10; Length 369;
Best Local Similarity 100.0%; Pred. No. 5.7; 0; Indels 0; Gaps 0;
Matches 22; Conservative 0; Mismatches 0

QY 1 CTGCAAGGGGTTTGTCTGGCG 22
DB 150 CTGCAAGGGGTTTGTCTGGCG 129

RESULT 11
US-09-534-846B-34372/c
Sequence 34372, Application US/09534846B
GENERAL INFORMATION:
APPLICANT: Seilhamer, Jeffrey J.
APPLICANT: Delegeane, Angelo M.
APPLICANT: Stuart, Susan G.
APPLICANT: Stuve, Laura L.
APPLICANT: Mullahy, Sara J.
APPLICANT: Naughton, Rebecca E.
TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING OR REGULATING GROWTH, DEVELOPMENT, AND
TITLE OF INVENTION: DIFFERENTIATION MOLECULES
FILE REFERENCE: PD-1021 CIP
CURRENT APPLICATION NUMBER: US/09/534,846B
CURRENT FILING DATE: 2002-01-09
Prior application data removed - refer to file wrapper or PALM
NUMBER OF SEQ ID NOS: 38710
SOFTWARE: PERL Program
SEQ ID NO 34372
LENGTH: 369
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
OTHER INFORMATION: Incyte ID No: hu00797632
NAME/KEY: unsure
LOCATION: 2, 78, 92, 175, 210, 271, 341
OTHER INFORMATION: a, t, c, g, or other
US-09-534-846B-34372

Query Match 100.0%; Score 22; DB 20; Length 369;
Best Local Similarity 100.0%; Pred. No. 5.7; 0; Indels 0; Gaps 0;
Matches 22; Conservative 0; Mismatches 0

QY 1 CTGCAAGGGGTTTGTCTGGCG 22
DB 150 CTGCAAGGGGTTTGTCTGGCG 129

RESULT 12
US-09-076-897-1393/c
Sequence 1393, Application US/09076897
GENERAL INFORMATION:
APPLICANT: Rosen, et. al.
TITLE OF INVENTION: Human Genes, Sequences, and Expression Products - P032
NUMBER OF SEQUENCES: 7180
CORRESPONDENCE ADDRESS:
ADDRESSEE: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/076,897
FILING DATE: May 13, 1998
CLASSIFICATION: 536
PRIOR APPLICATION DATA:

APPLICATION NUMBER: 60/047,005
FILING DATE: May 13, 1997
ATTORNEY/AGENT INFORMATION:
NAME: A. Anders Brookes
REGISTRATION NUMBER: 36,373
REFERENCE/DOCKET NUMBER: PO-32
TELEPHONE: (301) 309-8504
TELEFAX: (301) 309-8439
INFORMATION FOR SEQ ID NO: 1393:
SEQUENCE CHARACTERISTICS:
LENGTH: 404 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
US-09-076-897A-1393

Query Match 100.0%; Score 22; DB 14; Length 404;
Best Local Similarity 100.0%; Pred. No. 5.8;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTGCAAGGGGTTTGTCTGGCG 22
|||||

Db 159 CTGCAAGGGGTTTGTCTGGCG 138
|||||

RESULT 13

US-09-076-897A-1393/c
Sequence 1393, Application US/09076897A
GENERAL INFORMATION:
APPLICANT: Rosen, et. al.
TITLE OF INVENTION: Human Genes, Sequences, and Expression Products - P032
NUMBER OF SEQUENCES: 7180
CORRESPONDENCE ADDRESS:
ADDRESSEE: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: USA
ZIP: 20850

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/076,897A
FILING DATE: May 13, 1998
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: A. Anders Brookes
REGISTRATION NUMBER: 36,373
REFERENCE/DOCKET NUMBER: PO-32
TELEPHONE: (301) 309-8504
TELEFAX: (301) 309-8439
INFORMATION FOR SEQ ID NO: 1393:
SEQUENCE CHARACTERISTICS:
LENGTH: 404 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
US-09-076-897A-1393

Query Match 100.0%; Score 22; DB 14; Length 404;
Best Local Similarity 100.0%; Pred. No. 5.8;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTGCAAGGGGTTTGTCTGGCG 22
|||||

Db 159 CTGCAAGGGGTTTGTCTGGCG 138
|||||

RESULT 14

US-09-076-897B-1393/c
Sequence 1393, Application US/09076897B
GENERAL INFORMATION:
APPLICANT: Rosen, et. al.
TITLE OF INVENTION: Human Genes, Sequences, and Expression Products - P032
NUMBER OF SEQUENCES: 7180
CORRESPONDENCE ADDRESS:
ADDRESSEE: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: USA
ZIP: 20850

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/076,897B
FILING DATE: May 13, 1998
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: A. Anders Brookes
REGISTRATION NUMBER: 36,373
REFERENCE/DOCKET NUMBER: PO-32
TELEPHONE: (301) 309-8504
TELEFAX: (301) 309-8439
INFORMATION FOR SEQ ID NO: 1393:
SEQUENCE CHARACTERISTICS:
LENGTH: 404 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
US-09-076-897B-1393

Query Match 100.0%; Score 22; DB 14; Length 404;
Best Local Similarity 100.0%; Pred. No. 5.8;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTGCAAGGGGTTTGTCTGGCG 22
|||||

Db 159 CTGCAAGGGGTTTGTCTGGCG 138
|||||

RESULT 15

US-09-076-897C-1393/c
Sequence 1393, Application US/09076897C
GENERAL INFORMATION:
APPLICANT: Rosen, Craig, et al.
TITLE OF INVENTION: Human Genes, Sequences, and Expression Products 32
FILE REFERENCE: PO-32
CURRENT APPLICATION NUMBER: US/09/076,897C
CURRENT FILING DATE: 1998-05-13
PRIOR APPLICATION NUMBER: 1997-05-13
PRIOR FILING DATE: 60/047,005
NUMBER OF SEQ ID NOS: 7180
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 1393
LENGTH: 404
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
LOCATION: (122)

Mon Apr 21 10:29:47 2003

```

; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (123)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (124)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (129)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (290)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (331)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (333)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (341)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (365)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (392)
; OTHER INFORMATION: n equals a,t,g, or c
; US-09-076-897C-1393

```

```

Query Match      100.0%; Score 22; DB 14; Length 404;
Best Local Similarity 100.0%; Pred. No. 5.8;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 CTGCAAGGGGTTTGTCTGGGCG 22
    |||||
DB 159 CTGCAAGGGGTTTGTCTGGGCG 138

```

Search completed: April 18, 2003, 09:35:41
Job time : 1212 secs

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 18, 2003, 06:35:28; Search time 245,167 Seconds
(without alignments)
400.770 Million cell updates/sec

Title: US-09-270-437D-10

Perfect score: 22

Sequence: 1 CTGCAAGGGGTTTGTCTGGCG 22

Scoring table: IDENTITY NUC

Gapop 10.0, Gapext 1.0

Searched: 5897297 seqs, 2233080881 residues

Total number of hits satisfying chosen parameters: 11794594

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- Pending Patents NA New:*
- 1: /cgn2_6/ptodata/1/pna/PCT_NEW_COMB.seq:*
 - 2: /cgn2_6/ptodata/1/pna/US06_NEW_COMB.seq:*
 - 3: /cgn2_6/ptodata/1/pna/US07_NEW_COMB.seq:*
 - 4: /cgn2_6/ptodata/1/pna/US08_NEW_COMB.seq:*
 - 5: /cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:*
 - 6: /cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:*
 - 7: /cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:*
 - 8: /cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:*
 - 9: /cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:*
 - 10: /cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:*
 - 11: /cgn2_6/ptodata/1/pna/US60_NEW_COMB.seq:*
 - 12: /cgn2_6/ptodata/1/pna/US60_NEW_COMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	22	100.0	22	5	US-09-270-437D-10
2	22	100.0	369	9	Sequence 10, Appl
3	22	100.0	404	6	Sequence 34372, A
4	22	100.0	1740	9	Sequence 56784, A
5	22	100.0	1740	9	Sequence 347, App
6	22	100.0	1740	9	Sequence 478, App
7	22	100.0	1743	9	Sequence 483, App
8	22	100.0	1743	9	Sequence 447, App
9	22	100.0	1764	9	Sequence 450, App
10	22	100.0	1739	9	Sequence 428, App
11	22	100.0	2137	6	Sequence 485, App
12	22	100.0	2137	6	Sequence 44307, A
13	22	100.0	3897	8	Sequence 4307, A
14	22	100.0	4060	6	Sequence 4325, A
15	22	100.0	4060	6	Sequence 44305, A
16	22	100.0	4159	5	Sequence 44305, A
17	22	100.0	4181	9	Sequence 44305, A
18	22	100.0	4181	9	Sequence 44305, A
19	22	100.0	34555	9	Sequence 175, App
20	17.4	79.1	177251	7	Sequence 55, Appl
21	17.2	78.2	601	7	Sequence 479, App
22	17.2	78.2	1663	1	Sequence 15841, A
					Sequence 7089, Ap
					Sequence 3190, Ap

23	17.2	78.2	1795	9	US-10-144-771-2615
24	17.2	78.2	3730	11	US-60-453-135-1
25	17.2	78.2	3730	11	US-60-453-050-1
26	17.2	78.2	324604	1	PCT-US02-14597-56
27	17.2	78.2	373882	7	US-09-947-911-356
28	16.8	76.4	395	6	US-09-615-606A-15865
29	16.8	76.4	601	7	US-09-949-016-91995
30	16.8	76.4	601	7	US-09-949-016-91996
31	16.8	76.4	45755	7	US-09-949-016-15889
32	16.8	76.4	254405	7	US-09-949-016-14381
33	16.8	76.4	713680	7	US-09-949-016-14381
34	16.4	74.5	78125	7	US-09-949-016-16006
35	16.4	74.5	178256	6	US-09-948-124-62
36	16.4	74.5	15105345	5	US-09-948-128-55
37	16.2	73.6	285	6	US-09-531-113-28495
38	16.2	73.6	449	6	US-09-912-293-146413
39	16.2	73.6	500	8	US-10-375-582-237
40	16.2	73.6	553	9	US-10-314-344-11271
41	16.2	73.6	563	6	US-09-912-293-90564
42	16.2	73.6	859	6	US-09-699-652A-7
43	16.2	73.6	1226	9	US-10-144-771-18085
44	16.2	73.6	1560	9	US-10-282-122A-38395
45	16.2	73.6	2208	8	US-10-170-235-11352

ALIGNMENTS

RESULT 1

US-09-270-437D-10

; Sequence 10, Application US/09270437D

; GENERAL INFORMATION:

; APPLICANT: Chen, Yao-Tseng

; APPLICANT: Gure, Ali

; APPLICANT: Teang, Solam

; APPLICANT: Stockert, Elisabeth

; APPLICANT: Jager, Elke

; APPLICANT: Knuth, Alexander

; APPLICANT: Old, Lloyd J.

; TITLE OF INVENTION: Isolated Nucleic Acid Molecules Encoding Cancer Associated Antige

; FILE OF INVENTION: Antigens Per Se, And Uses Thereof

; FILE REFERENCE: LUD 5538.1

; CURRENT APPLICATION NUMBER: US/09/270,437D

; PRIOR FILING DATE: 1999-03-16

; PRIOR APPLICATION NUMBER: 09/061,709

; NUMBER OF SEQ ID NOS: 23

; SEQ ID NO 10

; LENGTH: 22

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

US-09-270-437D-10

Query Match

Best Local Similarity 100.0%; Score 22; DB 5; Length 22;

Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTGCAAGGGGTTTGTCTGGCG 22

Db 1 CTGCAAGGGGTTTGTCTGGCG 22

RESULT 2

US-10-314-344-34372/c

; Sequence 34372, Application US/10314344

; GENERAL INFORMATION:

; APPLICANT: Seilhamer, Jeffrey J.

; APPLICANT: Delegeane, Angelo M.

; APPLICANT: Stuart, Susan G.

; APPLICANT: Stuve, Laura L.

; APPLICANT: Mullaby, Sara J.

; APPLICANT: Naughton, Rebecca E.

us-09-270-437d-10.rnnpn

Mon Apr 21 10:29:50 2003

;; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING OR REGULATING GROWTH, DEVELOPMENT, AND

;; FILE REFERENCE: PD-1021-2 CON

;; CURRENT FILING DATE: 2002-12-05

;; PRIOR APPLICATION NUMBER: US/10/314,344

;; PRIOR FILING DATE: 2000-03-24

;; PRIOR APPLICATION NUMBER: 09/534,846

;; PRIOR FILING DATE: 1992-07-17

;; PRIOR APPLICATION NUMBER: 07/916,491

;; PRIOR FILING DATE: 1992-11-19

;; PRIOR APPLICATION NUMBER: 08/100,523

;; PRIOR FILING DATE: 1993-08-03

;; PRIOR APPLICATION NUMBER: 09/008,119

;; PRIOR FILING DATE: 1998-01-16

;; PRIOR APPLICATION NUMBER: 08/196,364

;; PRIOR FILING DATE: 1994-02-14

;; PRIOR APPLICATION NUMBER: 08/282,991

;; PRIOR FILING DATE: 1994-07-28

;; PRIOR APPLICATION NUMBER: 08/438,571

;; PRIOR FILING DATE: 1995-05-10

;; PRIOR APPLICATION NUMBER: 08/179,873

;; PRIOR FILING DATE: 1994-01-11

;; PRIOR APPLICATION NUMBER: 08/504,732

;; PRIOR FILING DATE: 1995-07-20

;; Remaining Prior Application data removed - See File Wrapper or PALM.

;; NUMBER OF SEQ ID NOS: 38710

;; SOFTWARE: PERL Program

;; SEQ ID NO 34372

;; LENGTH: 369

;; TYPE: DNA

;; ORGANISM: Homo sapiens

;; NAME/KEY: misc feature

;; LOCATION: 2, 78, 92, 175, 210, 271, 341

;; OTHER INFORMATION: a, t, c, g, or other

US-10-314-344-34372

Query Match 100.0%; Score 22; DB 9; Length 369;

Best Local Similarity 100.0%; Pred. No. 1.1; Indels 0; Gaps 0;

Matches 22; Conservative 0; Mismatches 0;

Qy 1 CTGCAAGGGGTTTCTGGGCG 22

Db 150 CTGCAAGGGGTTTCTGGGCG 129

RESULT 3

US-09-912-293-56784/c

;; Sequence 56784, Application US/09912293

;; GENERAL INFORMATION:

;; APPLICANT: Rosen, et. al.

;; TITLE OF INVENTION: Human Genes, Sequences, and Expression Products 100

;; FILE REFERENCE: PO-100

;; CURRENT FILING DATE: 2001-07-26

;; PRIOR APPLICATION NUMBER: US/09/912,293

;; PRIOR FILING DATE: 2001-07-26

;; PRIOR APPLICATION NUMBER: 08/103,744

;; PRIOR FILING DATE: 1993-08-09

;; PRIOR APPLICATION NUMBER: 09/249,651

;; PRIOR FILING DATE: 1999-02-12

;; PRIOR APPLICATION NUMBER: 08/104,507

;; PRIOR FILING DATE: 1993-08-09

;; PRIOR APPLICATION NUMBER: 08/196,363

;; PRIOR FILING DATE: 1994-02-15

;; PRIOR APPLICATION NUMBER: 09/859,490

;; PRIOR FILING DATE: 2001-05-18

;; PRIOR APPLICATION NUMBER: 08/196,362

;; PRIOR FILING DATE: 1994-02-15

;; PRIOR APPLICATION NUMBER: 08/221,623

;; PRIOR FILING DATE: 1994-03-31

;; PRIOR APPLICATION NUMBER: 08/220,691

;; PRIOR FILING DATE: 1994-03-31

;; PRIOR APPLICATION NUMBER: 09/741,830

;; PRIOR FILING DATE: 2000-12-22

;; PRIOR APPLICATION NUMBER: 09/813,155

;; PRIOR FILING DATE: 2001-03-21

;; Remaining Prior Application data removed - See File Wrapper or PALM.

;; NUMBER OF SEQ ID NOS: 244538

;; SEQ ID NO 56784

;; LENGTH: 404

;; TYPE: DNA

;; ORGANISM: Homo sapiens

;; NAME/KEY: misc feature

;; LOCATION: (122)..(124)

;; OTHER INFORMATION: n is equal to a,t,g, or c

;; FEATURE:

;; NAME/KEY: misc feature

;; LOCATION: (129)..(129)

;; OTHER INFORMATION: n is equal to a,t,g, or c

;; FEATURE:

;; NAME/KEY: misc feature

;; LOCATION: (290)..(290)

;; OTHER INFORMATION: n is equal to a,t,g, or c

;; FEATURE:

;; NAME/KEY: misc feature

;; LOCATION: (331)..(331)

;; OTHER INFORMATION: n is equal to a,t,g, or c

;; FEATURE:

;; NAME/KEY: misc feature

;; LOCATION: (333)..(333)

;; OTHER INFORMATION: n is equal to a,t,g, or c

;; FEATURE:

;; NAME/KEY: misc feature

;; LOCATION: (341)..(341)

;; OTHER INFORMATION: n is equal to a,t,g, or c

;; FEATURE:

;; NAME/KEY: misc feature

;; LOCATION: (365)..(365)

;; OTHER INFORMATION: n is equal to a,t,g, or c

;; FEATURE:

;; NAME/KEY: misc feature

;; LOCATION: (392)..(392)

;; OTHER INFORMATION: n is equal to a,t,g, or c

US-09-912-293-56784

Query Match 100.0%; Score 22; DB 6; Length 404;

Best Local Similarity 100.0%; Pred. No. 1.1;

Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTGCAAGGGGTTTCTGGGCG 22

Db 159 CTGCAAGGGGTTTCTGGGCG 138

RESULT 4

US-10-313-986-347/c

;; Sequence 347, Application US/10313986

;; GENERAL INFORMATION:

;; APPLICANT: Foy, Teresa M.

;; APPLICANT: McNabb, Andria

;; APPLICANT: Watanabe, Yoshihiro

;; APPLICANT: Reed, Steven G.

;; APPLICANT: Wang, Tongcong

;; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY

;; FILE REFERENCE: 210121.455C19

;; CURRENT FILING DATE: 2002-12-04

;; NUMBER OF SEQ ID NOS: 560

;; SOFTWARE: FastSeq for Windows Version 4.0

;; SEQ ID NO 347

;; LENGTH: 1740

; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-313-986-347

Query Match 100.0%; Score 22; DB 9; Length 1740;
Best Local Similarity 100.0%; Pred. No. 1.3;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTGCAAGGGGTTTGTCTGGCG 22
|||||
DB 498 CTGCAAGGGGTTTGTCTGGCG 477

RESULT 5

US-10-313-986-478/c
; Sequence 478, Application US/10313986
; GENERAL INFORMATION:
; APPLICANT: Foy, Teresa M.
; APPLICANT: McNabb, Andria
; APPLICANT: Watanabe, Yoshihiro
; APPLICANT: Reed, Steven G.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.455C19
; CURRENT APPLICATION NUMBER: US/10/313,986
; CURRENT FILING DATE: 2002-12-04
; NUMBER OF SEQ ID NOS: 560
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 478
; LENGTH: 1740
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-313-986-478

Query Match 100.0%; Score 22; DB 9; Length 1740;
Best Local Similarity 100.0%; Pred. No. 1.3;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTGCAAGGGGTTTGTCTGGCG 22
|||||
DB 498 CTGCAAGGGGTTTGTCTGGCG 477

RESULT 6

US-10-313-986-483/c
; Sequence 483, Application US/10313986
; GENERAL INFORMATION:
; APPLICANT: Foy, Teresa M.
; APPLICANT: McNabb, Andria
; APPLICANT: Watanabe, Yoshihiro
; APPLICANT: Reed, Steven G.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.455C19
; CURRENT APPLICATION NUMBER: US/10/313,986
; CURRENT FILING DATE: 2002-12-04
; NUMBER OF SEQ ID NOS: 560
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 483
; LENGTH: 1740
; TYPE: DNA
; ORGANISM: Primate
US-10-313-986-483

Query Match 100.0%; Score 22; DB 9; Length 1740;
Best Local Similarity 100.0%; Pred. No. 1.3;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTGCAAGGGGTTTGTCTGGCG 22
|||||
DB 498 CTGCAAGGGGTTTGTCTGGCG 477

RESULT 7

US-10-313-986-447/c
; Sequence 447, Application US/10313986
; GENERAL INFORMATION:
; APPLICANT: Foy, Teresa M.
; APPLICANT: McNabb, Andria
; APPLICANT: Watanabe, Yoshihiro
; APPLICANT: Reed, Steven G.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.455C19
; CURRENT APPLICATION NUMBER: US/10/313,986
; CURRENT FILING DATE: 2002-12-04
; NUMBER OF SEQ ID NOS: 560
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 447
; LENGTH: 1743
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-313-986-447

Query Match 100.0%; Score 22; DB 9; Length 1743;
Best Local Similarity 100.0%; Pred. No. 1.3;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTGCAAGGGGTTTGTCTGGCG 22
|||||
DB 498 CTGCAAGGGGTTTGTCTGGCG 477

RESULT 8

US-10-313-986-450/c
; Sequence 450, Application US/10313986
; GENERAL INFORMATION:
; APPLICANT: Foy, Teresa M.
; APPLICANT: McNabb, Andria
; APPLICANT: Watanabe, Yoshihiro
; APPLICANT: Reed, Steven G.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.455C19
; CURRENT APPLICATION NUMBER: US/10/313,986
; CURRENT FILING DATE: 2002-12-04
; NUMBER OF SEQ ID NOS: 560
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 450
; LENGTH: 1743
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-313-986-450

Query Match 100.0%; Score 22; DB 9; Length 1743;
Best Local Similarity 100.0%; Pred. No. 1.3;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTGCAAGGGGTTTGTCTGGCG 22
|||||
DB 498 CTGCAAGGGGTTTGTCTGGCG 477

RESULT 9

US-10-313-986-428/c
; Sequence 428, Application US/10313986
; GENERAL INFORMATION:
; APPLICANT: Foy, Teresa M.
; APPLICANT: McNabb, Andria
; APPLICANT: Watanabe, Yoshihiro
; APPLICANT: Reed, Steven G.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.455C19
; CURRENT APPLICATION NUMBER: US/10/313,986
; CURRENT FILING DATE: 2002-12-04
; NUMBER OF SEQ ID NOS: 560
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 428
; LENGTH: 1740
; TYPE: DNA
; ORGANISM: Primate
US-10-313-986-428

Mon Apr 21 10:29:50 2003

;; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
;; FILE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER

;; FILE REFERENCE: 210121.455C19
;; CURRENT APPLICATION NUMBER: US/10/313,986
;; CURRENT FILING DATE: 2002-12-04
;; NUMBER OF SEQ ID NOS: 560
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 428
;; LENGTH: 1764
;; TYPE: DNA
;; ORGANISM: Homo sapiens
US-10-313-986-428

Query Match 100.0%; Score 22; DB 9; Length 1764;
Best Local Similarity 100.0%; Pred. No. 1.3;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTGCAAGGGGTTTCTGGGCG 22
|||
DB 519 CTGCAAGGGGTTTCTGGGCG 498

RESULT 10
US-10-313-986-485/c

;; Sequence 485, Application US/10313986
;; GENERAL INFORMATION:
;; APPLICANT: Poy, Teresa M.
;; APPLICANT: McNabb, Andria
;; APPLICANT: Watanabe, Yoshihiro
;; APPLICANT: Reed, Steven G.
;; APPLICANT: Wang, Tongtong
;; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
;; FILE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
;; FILE REFERENCE: 210121.455C19
;; CURRENT APPLICATION NUMBER: US/10/313,986
;; CURRENT FILING DATE: 2002-12-04
;; NUMBER OF SEQ ID NOS: 560
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 485
;; LENGTH: 1799
;; TYPE: DNA
;; ORGANISM: Homo sapiens
US-10-313-986-485

Query Match 100.0%; Score 22; DB 9; Length 1799;
Best Local Similarity 100.0%; Pred. No. 1.3;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTGCAAGGGGTTTCTGGGCG 22
|||
DB 517 CTGCAAGGGGTTTCTGGGCG 496

RESULT 11
US-09-724-676-44307/c

;; Sequence 44307, Application US/09724676
;; GENERAL INFORMATION:
;; APPLICANT: Compugen LTD
;; TITLE OF INVENTION: Variants of alternative splicing
;; FILE REFERENCE: 129181.4 Compugen
;; CURRENT APPLICATION NUMBER: US/09/724,676
;; CURRENT FILING DATE: 2000-11-28
;; NUMBER OF SEQ ID NOS: 97222
;; SOFTWARE: PatentIn version 3.2
;; SEQ ID NO 44307
;; LENGTH: 2137
;; TYPE: DNA
;; ORGANISM: Homo sapiens
US-09-724-676-44307

Query Match 100.0%; Score 22; DB 6; Length 2137;
Best Local Similarity 100.0%; Pred. No. 1.4;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTGCAAGGGGTTTCTGGGCG 22
|||
DB 764 CTGCAAGGGGTTTCTGGGCG 743

RESULT 12
US-09-724-676A-44307/c

;; Sequence 44307, Application US/09724676A
;; GENERAL INFORMATION:
;; APPLICANT: Compugen LTD
;; TITLE OF INVENTION: Variants of alternative splicing
;; FILE REFERENCE: 129181.4 Compugen
;; CURRENT APPLICATION NUMBER: US/09/724,676A
;; CURRENT FILING DATE: 2000-11-28
;; NUMBER OF SEQ ID NOS: 97222
;; SOFTWARE: PatentIn version 3.2
;; SEQ ID NO 44307
;; LENGTH: 2137
;; TYPE: DNA
;; ORGANISM: Homo sapiens
US-09-724-676A-44307

Query Match 100.0%; Score 22; DB 6; Length 2137;
Best Local Similarity 100.0%; Pred. No. 1.4;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTGCAAGGGGTTTCTGGGCG 22
|||
DB 764 CTGCAAGGGGTTTCTGGGCG 743

RESULT 13
US-10-170-235-31225/c

;; Sequence 31225, Application US/10170235
;; GENERAL INFORMATION:
;; APPLICANT: VENTER, J. Craig
;; TITLE OF INVENTION: KITS, SUCH AS NUCLEIC ACID ARRAYS, COMPRISING A MAJORITY OF HUMAN
;; FILE OF INVENTION: TRANSCRIPTS, FOR DETECTING EXPRESSION AND OTHER USES THEREOF
;; FILE REFERENCE: CL001380
;; CURRENT APPLICATION NUMBER: US/10/170,235
;; CURRENT FILING DATE: 2003-03-17
;; NUMBER OF SEQ ID NOS: 42514
;; SEQ ID NO 31225
;; LENGTH: 3897
;; TYPE: DNA
;; ORGANISM: HUMAN
US-10-170-235-31225

Query Match 100.0%; Score 22; DB 8; Length 3897;
Best Local Similarity 100.0%; Pred. No. 1.5;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTGCAAGGGGTTTCTGGGCG 22
|||
DB 334 CTGCAAGGGGTTTCTGGGCG 313

RESULT 14
US-09-724-676-44305/c

;; Sequence 44305, Application US/09724676
;; GENERAL INFORMATION:
;; APPLICANT: Compugen LTD
;; TITLE OF INVENTION: Variants of alternative splicing
;; FILE REFERENCE: 129181.4 Compugen
;; CURRENT APPLICATION NUMBER: US/09/724,676
;; CURRENT FILING DATE: 2000-11-28
;; NUMBER OF SEQ ID NOS: 97222
;; SOFTWARE: PatentIn version 3.2
;; SEQ ID NO 44305
;; LENGTH: 4060
;; TYPE: DNA
;; ORGANISM: Homo sapiens

US-09-724-676-44305

Query Match 100.0%; Score 22; DB 6; Length 4060;
Best Local Similarity 100.0%; Pred. No. 1.5;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTGCAAGGGGTTTGTCTGGCG 22
DB 764 CTGCAAGGGGTTTGTCTGGCG 743

RESULT 15

US-09-724-676A-44305/c
; Sequence 44305, Application US/09724676A
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676A
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 44305
; LENGTH: 4060
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-724-676A-44305

Query Match 100.0%; Score 22; DB 6; Length 4060;
Best Local Similarity 100.0%; Pred. No. 1.5;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTGCAAGGGGTTTGTCTGGCG 22
DB 764 CTGCAAGGGGTTTGTCTGGCG 743

Search completed: April 18, 2003, 10:03:47
Job time : 256.167 secs

GenCore version 5.1.4 p5 4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 18, 2003, 04:47:40 ; Search time 766 Seconds
(without alignments)
465.145 Million cell updates/sec

Title: US-09-270-437D-10

Perfect score: 22
Sequence: 1 ctgcaagggttttctgtggcg 22

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST.*

1: em_estba.*

2: em_esthum.*

3: em_estm.*

4: em_estm.*

5: em_estov.*

6: em_estpl.*

7: em_estro.*

8: em_hic.*

9: gb_est1.*

10: gb_est2.*

11: gb_hic.*

12: gb_est3.*

13: gb_est4.*

14: gb_est5.*

15: em_estfun.*

16: em_estom.*

17: gb_gss.*

18: em_gss_hum.*

19: em_gss_inv.*

20: em_gss_pln.*

21: em_gss_vrt.*

22: em_gss_fun.*

23: em_gss_nam.*

24: em_gss_mus.*

25: em_gss_other.*

26: em_gss_pro.*

27: em_gss_rod.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	22	100.0	790	10	BE545535
C 2	20.4	92.7	743	13	BI520069
C 3	19.4	88.2	537	17	AZ019431
C 4	19	86.4	813	12	BG699047
C 5	18.8	85.5	372	17	AQ022143
C 6	18.8	85.5	466	9	AL597825

7	18	81.8	615	17	DR6F10S	AL744135
8	17.8	80.9	375	14	BQ166727	Danio rer
9	17.8	80.9	714	14	BM728827	BQ166727 WHE0924 D
10	17.8	80.9	719	17	BH592483	UI-E-EOI-
11	17.8	80.9	1253	12	BF183419	BH592483 BOG090TF
12	17.4	79.1	271	9	A1545505	BF183419 601809561
13	17.4	79.1	543	13	BI338062	A1545505 fb66d03.x
14	17.4	79.1	556	12	BF230968	BI338062 361720 MA
15	17.4	79.1	573	12	BE899763	BF230968 252845 BA
16	17.4	79.1	575	9	A1658303	BE899763 202763 MA
17	17.4	79.1	583	13	BI536530	A1658303 fc21c10.y
18	17.4	79.1	689	12	BG027293	BI536530 393570 MA
19	17.4	79.1	891	12	BE972093	BG027293 602296004
20	17.4	79.1	904	14	BQ715397	BE972093 601651550
21	17.4	79.1	952	12	BG106781	BQ715397 AGENCOURT
22	17.4	79.1	1226	13	BM563299	BG106781 602290632
23	17.2	78.2	267	17	AG063729	BM563299 AGENCOURT
24	17.2	78.2	289	10	BB231263	AG063729 Pan trogl
25	17.2	78.2	332	14	W10000	BB231263 BB291263
26	17.2	78.2	335	9	AA200129	W10000 mas1f08.r1
27	17.2	78.2	342	12	BF919964	AA200129 mu10a06.x
28	17.2	78.2	344	14	W47750	BF919964 MR1-NT017
29	17.2	78.2	379	9	A1159434	W47750 mc84d01.r1
30	17.2	78.2	395	17	AQ340144	A1159434 VZ79b07.r
31	17.2	78.2	397	9	AA270892	AQ340144 HS_5021.A
32	17.2	78.2	400	9	AI745603	AA270892 VZ79b08.r
33	17.2	78.2	404	10	AW912298	AI745603 WC33F02.x
34	17.2	78.2	415	12	BF899524	AW912298 uf43h12.y
35	17.2	78.2	416	14	W54455	BF899524 QV1-MT022
36	17.2	78.2	420	9	AA880520	W54455 md08b09.r1
37	17.2	78.2	425	9	AI836236	AA880520 VV91902.r
38	17.2	78.2	425	10	BB777528	AI836236 UI-M-AP0-
39	17.2	78.2	427	10	BB778560	BB777528 BB777528
40	17.2	78.2	429	14	W20987	BB778560 BB788560
41	17.2	78.2	431	10	AW492429	W20987 mb89e12.r1
42	17.2	78.2	434	9	AI154703	AW492429 UI-N-BH3-
43	17.2	78.2	436	10	BB711444	AI154703 ud35907.r
44	17.2	78.2	436	17	AQ313390	BB711444 BB711444
45	17.2	78.2	442	17	AQ674038	AQ313390 RPC111-11
						AQ674038 HS_5477_B

ALIGNMENTS

RESULT 1
BE545535/c
LOCUS BE545535 790 bp mRNA linear EST 09-AUG-2000
DEFINITION 601070391F1 NIH_MGC_12 Homo sapiens cDNA clone IMAGE:3456407 5', mRNA sequence.
ACCESSION BE545535
VERSION BE545535
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 790)
AUTHORS NIH-MGC <http://mgi.nci.nih.gov/>
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: Incyte Genomics, Inc.
Clone Sequencing by: Incyte Genomics, Inc.
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLAM8444 row: f column: 24
High quality sequence stop: 667.
FEATURES
Location/Qualifiers
1..790

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:3456407"
/clone_lib="NIH_MGC_12"
/tissue_type="cervical carcinoma cell line"
/lab_host="DH10B"
/notes="Organ: cervix; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 1.4 kb. Library prepared by Life
Technologies."
BASE COUNT      221 a 193 c 209 g 167 t
ORIGIN
Query Match      100.0%; Score 22; DB 10; Length 790;
Best Local Similarity 100.0%; Pred. No. 14;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTGCAAGGGGTTTGTCTGGGCG 22
|||||
Db 416 CTGCAAGGGGTTTGTCTGGGCG 395

RESULT 2
BI520069/c
LOCUS      743 bp mRNA linear EST 29-AUG-2001
DEFINITION 603071386F1 NIH_MGC_119 Homo sapiens cDNA clone IMAGE:5163429 5',
RNA sequence.
ACCESSION  BI520069.1 GI:15344861
VERSION     BI520069.1
KEYWORDS    EST.
SOURCE      human.
ORGANISM    Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 743)
AUTHORS     NIH-MGC http://mgc.nci.nih.gov/
TITLE       National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL     Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-x@mail.nih.gov
            Tissue Procurement: Life Technologies, Inc.
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone Distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: L14M11405 row: 1 column: 22
            High quality sequence start: 9
            High quality sequence stop: 704.

FEATURES
source
1..743
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5163429"
/clone_lib="NIH_MGC_119"
/tissue_type="medulla"
/lab_host="DH10B"
/notes="Organ: brain; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: EcoRV (destroyed); RNA source normal medulla from
anonymous male age 27. Library is oligo-dt primed and
directionally cloned (EcoRV site is destroyed upon
cloning). Average insert size 1.3 kb, insert size range
0.9-3 kb. Library is normalized and enriched for
full-length clones and was constructed by C. Gruber
(invitrogen). Research Genetics tracking code 013. Note:
this is a NIH_MGC Library."
BASE COUNT      226 a 167 c 174 g 176 t
ORIGIN
Query Match      92.7%; Score 20.4; DB 13; Length 743;
Best Local Similarity 95.5%; Pred. No. 72;
Matches 21; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:3456407"
/clone_lib="NIH_MGC_12"
/tissue_type="cervical carcinoma cell line"
/lab_host="DH10B"
/notes="Organ: cervix; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 1.4 kb. Library prepared by Life
Technologies."
BASE COUNT      221 a 193 c 209 g 167 t
ORIGIN
Query Match      100.0%; Score 22; DB 10; Length 790;
Best Local Similarity 100.0%; Pred. No. 14;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTGCAAGGGGTTTGTCTGGGCG 22
|||||
Db 416 CTGCAAGGGGTTTGTCTGGGCG 395

RESULT 3
AZ019431
LOCUS      537 bp DNA linear GSS 25-FEB-2000
DEFINITION RPCI-23-299D12.TJ RPCI-23 Mus musculus genomic clone RPCI-23-299D12
, DNA sequence.
ACCESSION  AZ019431
VERSION     AZ019431.1 GI:7094815
KEYWORDS    GSS.
SOURCE      house mouse.
ORGANISM    Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1 (bases 1 to 537)
AUTHORS     Zhao, S., Nierman, W., Feldblyum, T., Malek, J., Shatman, S., Akinret
            B., Levins, M., McGann, S., Tsegaye, G., Geer, K., Krol, M., de Jong, P.
            and Fraser, C.M.
            Mouse BAC End Sequences from Library RPCI-23
            Unpublished (1999)
            Contact: Shaying Zhao
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: szhao@tigr.org
            Clones are derived from the mouse BAC library RPCI-23. For BAC
            library availability, please contact Pieter de Jong
            (pieter@dejong.med.buffalo.edu). Clones may be purchased from
            BACPAC Resources (http://bacpac.med.buffalo.edu/orderingframe.htm)
            or from Resea ch Genetics (info@resgen.com). BAC end page:
            http://www.tigr.org/tdb/bac_ends/mouse/bac_end_intro.html
            Plate: 299 row: D column: 12
            Seq primer: SP6
            Class: BAC ends.
            Location/Qualifiers
            1..537
            /organism="Mus musculus"
            /strain="C57BL/6J"
            /db_xref="taxon:10090"
            /clone="RPCI-23-299D12"
            /clone_lib="RPCI-23"
            /sex="Female"
            /lab_host="DH10B"
            /note="Organ: Kidney/Brain; Vector: pBACe3.6; Site 1:
            EcoRI; Site 2: EcoRI; Female C57BL/6J mouse kidney and/or
            brain genomic DNA was isolated and partially digested
            with a combination of EcoRI and EcoRI Methylase. Size
            selected DNA was cloned into the pBACe3.6 vector at the
            EcoRI sites. The ligation products were transformed into
            DH10B electrocompetent cells (BRL Life Technologies)."
            109 a 135 c 112 g 180 t 1 others
BASE COUNT      109 a 135 c 112 g 180 t 1 others
ORIGIN
Query Match      88.2%; Score 19.4; DB 17; Length 537;
Best Local Similarity 95.2%; Pred. No. 1.8e+02;
Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 TGCAAGGGGTTTGTCTGGGCG 22
|||||
Db 426 TGCAAGGGGTTTGTCTGGGCG 446

RESULT 4
BG699047/c
LOCUS      813 bp mRNA linear EST 07-MAY-2001
DEFINITION 602678681F1 NIH_MGC_95 Homo sapiens cDNA clone IMAGE:4811420 5',
mRNA sequence.

```


IMAGE:3716937 3' similar to contains element MSRI1 repetitive element ;, mRNA sequence.

AI545505
AI545505.1 GI:4462878
EST.
zebrafish.

ORGANISM
Danio rerio
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes; Cyprinidae; Danio.

REFERENCE
1 (bases 1 to 271)
Clark, M., Johnson, S.L., Lehrach, H., Lee, R., Li, F., Marra, M., Eddy, S., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wyllie, T., Underwood, K., Steptoe, M., Theising, B., Allen, M., Bowers, Y., Person, B., Swaller, R., Gibbons, M., Pape, D., Harvey, N., Schurk, R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R., Waterston, R. and Wilson, R.

TITLE
WashU Zebrafish EST Project 1998

JOURNAL
Unpublished (1998)

COMMENT
Contact: Stephen L. Johnson
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: zbrafish@watson.wustl.edu
CDNA Library Preparation: Matthew Clark. CDNA Library Arrayed by: Matthew Clark. DNA Sequencing by: Washington University Genome Sequencing Center Clone Distribution: Genome Systems, St. Louis, Missouri (web address: www.genomesystems.com) (email contact: info@genomesystems.com) and Research Genetics, Huntsville, Alabama (web address: www.resgen.com) (email contact: info@resgen.com) and ResourceZentrumPrimarDatenbank, Berlin, Germany (web address: www.rzpd.de)
Seq primer: r7 ET from Amersham
High quality sequence stop: 269
POLYA=No.

FEATURES
source
1..271
/organism="Danio rerio"
/db_xref="taxon:7955"
/clone="IMAGE:3716937"
/clone_lib="Zebrafish WashU MPIMG EST"
/sex="mixed"
/tissue_type="26 somite embryos, adult livers, shield stage embryos"
/lab_host="XLI-blue MRF"
/note="Vector: pSPORI1; Site 1: NotI; Site 2: SalI; 1st strand cDNA was primed with a Not I - oligo(dT)15 primer [5'-pGACTGATTCACATCGGAGCGCCCTTTTTTTTTTTT3']; double-stranded cDNA was ligated to Sal I adaptors (BRL), digested with Not I and cloned into the Not I and Sal I sites of the pSPORI1 vector (BRL). Library was constructed by Matthew Clark (Lehrach lab; ICRF, London and Max Planck Institut fuer Molekulare Genetik, Berlin). cDNAs for EST analysis were selected following oligonucleotide hybridization fingerprinting of arrayed clones from zebrafish late somitogenesis (26 ss), adult liver or embryonic shield stage (5.6 h) libraries. Fingerprint data were used to computationally cluster cDNAs, and a single cDNA from each cluster was chosen for sequencing. In some cases multiple members of the same cluster were sequenced to assess clustering parameters or single clone control."

BASE COUNT
31 a 50 c 63 g 127 t

ORIGIN
79.1%; Score 17.4; DB 9; Length 271;
Query Match 94.7%; Pred. No. 1.2e+03;
Best Local Similarity 0; Mismatches 1; Indels 0; Gaps 0;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
4 CAAGGGGTTTTCGTGGCG 22

JOURNAL
COMMENT

gland cDNA library
Unpublished (2000)
Contact: Sonstegard TS
USDA, ARS, Beltsville Agricultural Research Center
Bldg. 200 Rm 2A, Beltsville, MD 20705, USA
Tel: 301 504 8416
Fax: 301 504 8414
Email: tads@anri.barc.usda.gov
Single pass sequencing. Bases called and alt trimmed with phred
v0.980904.e. Vector identified by cross_match with the -minscore 18
and -minmatch 12 options.
PCR Primers
FORWARD: AGGAACAGCTATGACCAT
BACKWARD: GTTTCCCGTACGACG
Plate: 108 row: H column: 3
Seq primer: ATTAGGTGACACTATAG.
Location/Qualifiers

FEATURES
source

1..556
/organism="Bos taurus"
/db_xref="taxon:9913"
/clone_lib="BARC SBOV"
/tissue_type="pooled"
/lab_host="DH10B"

/note="Vector: PCMV SPORT6; Site 1: NotI; Site 2: SalI;
Library made from pooled mRNA isolated from mammary
tissues at eight physiological, developmental, and disease
states."

BASE COUNT 166 a 94 c 111 g 185 t
ORIGIN

Query Match 79.1%; Score 17.4; DB 12; Length 556;
Best Local Similarity 94.7%; Pred. No. 1.5e+03;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 TGCAGGGGTTTGTCTGGG 20
|||||
Db 320 TGCAGGGGTTTGTCTGTG 338

RESULT 15

BE899763

LOCUS BE899763 573 bp mRNA linear EST 25-APR-2001
DEFINITION 202763 MARC 2BOV Bos taurus cDNA 5', mRNA sequence.
ACCESSION BE899763
VERSION BE899763.1 GI:10387237
KEYWORDS EST.
SOURCE COW.

ORGANISM

Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
Bovidae; Bovinae; Bos.

REFERENCE

AUTHORS

Smith, T.P.L., Grosse, W.M., Freking, B.A., Roberts, A.J., Stone, R.T.,
Cases, E., Wray, J.E., White, J., Cho, J., Fahrenkrug, S.C., Bennett,
G.L., Heaton, M.P., Laegreid, W.W., Rohrer, G.A., Chitko-McKown, C.G.,
Perteau, G., Holt, I., Karamycheva, S., Liang, F., Quackenbush, J. and
Keefe, J.W.

TITLE

Sequence evaluation of four pooled-tissue normalized bovine cDNA
libraries and construction of a gene index for cattle

Genome Res. 11 (4), 626-630 (2001)

21180013

JOURNAL

MEDLINE

COMMENT

Contact: Smith TPL
USDA, ARS, US Meat Animal Research Center
PO Box 166, Clay Center, NE 68933-0166, USA
Tel: 402 762 4366
Fax: 402 762 4390

Email: smith@mail.marc.usda.gov

Single pass sequencing. Bases called and alt trimmed with phred
v0.980904.e. Vector identified by cross_match with the -minscore 18
and -minmatch 12 options.

PCR Primers

FORWARD: AGGAACAGCTATGACCAT

BACKWARD: GTTTCCCGTACGACG

Plate: 42 row: G column: 17
Seq primer: ATTAGGTGACACTATAG.
Location/Qualifiers

FEATURES
source

1..573
/organism="Bos taurus"
/db_xref="taxon:9913"
/clone_lib="MARC 2BOV"
/tissue_type="pooled"
/lab_host="DH10B"

/note="Vector: PCMV SPORT6; Site 1: NotI; Site 2: SalI;
Library made from pooled tissue from testis, thymus,
semitendinosus muscle, longissimus muscle, pancreas,
adrenal, and endometrium."

165 a 97 c 112 g 199 t

BASE COUNT
ORIGIN

Query Match 79.1%; Score 17.4; DB 12; Length 573;
Best Local Similarity 94.7%; Pred. No. 1.5e+03;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 TGCAGGGGTTTGTCTGGG 20
|||||
Db 311 TGCAGGGGTTTGTCTGTG 329

Search completed: April 18, 2003, 07:31:49
Job time : 771.167 secs

GenCore version 5.1.4 p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 18, 2003, 04:46:26 ; Search time 292 Seconds
(without alignments)
2192.677 Million cell updates/sec

Title: US-09-270-437D-11
Perfect score: 22
Sequence: 1 tcttgcgcgcgtgcgcctcag 22

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl.:

1: gb.ba.*
2: gb.htg.*
3: gb.in.*
4: gb.om.*
5: gb.ov.*
6: gb.pat.*
7: gb.ph.*
8: gb.pl.*
9: gb.pr.*
10: gb.ro.*
11: gb.sts.*
12: gb.sy.*
13: gb.un.*
14: gb.vi.*
15: em.ba.*
16: em.fun.*
17: em.hum.*
18: em.in.*
19: em.mu.*
20: em.om.*
21: em.or.*
22: em.ov.*
23: em.pat.*
24: em.ph.*
25: em.pl.*
26: em.ro.*
27: em.sts.*
28: em.un.*
29: em.vi.*
30: em.htg.hum.*
31: em.htg.inv.*
32: em.htg.other.*
33: em.htg.mus.*
34: em.htg.pln.*
35: em.htg.rtd.*
36: em.htg.mam.*
37: em.htg.vrt.*
38: em.sy.*
39: em.htgo.hum.*
40: em.htgo.mus.*
41: em.htgo.other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	22	100.0	1946	6	AR171866	AR171866 Sequence
C 2	22	100.0	147767	2	AC105030	AC105030 Homo sapi
C 3	22	100.0	168613	9	AC091133	AC091133 Homo sapi
C 4	18.8	85.5	12110	1	AE001245	AE001245 Treponema
C 5	18.8	85.5	175585	2	AC118983	AC118983 Rattus no
C 6	17.8	80.9	2023	9	AF463492	AF463492 Homo sapi
C 7	17.8	80.9	2082	9	AK027725	AK027725 Homo sapi
C 8	17.8	80.9	2477	9	BC026308	BC026308 Homo sapi
C 9	17.8	80.9	146180	9	AC007193	AC007193 Homo sapi
C 10	17.8	80.9	210868	2	AC099392	AC099392 Rattus no
C 11	17.4	79.1	352	11	R10G148A	D14746 Oryza sativ
C 12	17.4	79.1	7419	9	AB061849	AB061849 Homo sapi
C 13	17.4	79.1	83052	2	CNS08C9Y	AL772420 Oryza sat
C 14	17.4	79.1	132592	2	AL589792	AL589792 Homo sapi
C 15	17.4	79.1	166518	9	HS120G22	AL031847 Human DNA
C 16	17.4	79.1	173627	9	AC004963	AC004963 Homo sapi
C 17	17.2	78.2	552	3	AF153349	AF153349 Leishmani
C 18	17.2	78.2	818	9	HS3336031	AF153349 Leishmani
C 19	17.2	78.2	888	9	HUM21SEQJ	AJ336031 Homo sapi
C 20	17.2	78.2	1426	3	AY069730	M97565 Human (Clon
C 21	17.2	78.2	1686	10	BC030347	AY069730 Drosophill
C 22	17.2	78.2	1727	10	AF214013	BC030347 Mus muscu
C 23	17.2	78.2	2719	1	AF051693	AF214013 Mus muscu
C 24	17.2	78.2	10757	1	AE004561	AF051693 Pseudomon
C 25	17.2	78.2	10981	1	AE005890	AE004561 Pseudomon
C 26	17.2	78.2	15000	3	LMFL2267	AE005890 Caulobact
C 27	17.2	78.2	15267	3	AB055099	AL357593 Leishmani
C 28	17.2	78.2	22990	3	AB055101	AB055099 Drosophill
C 29	17.2	78.2	23007	3	AB055100	AB055101 Drosophill
C 30	17.2	78.2	23244	3	AB055098	AB055100 Drosophill
C 31	17.2	78.2	61204	2	AC017132	AB055098 Drosophill
C 32	17.2	78.2	74169	2	AC128268	AC017132 Drosophill
C 33	17.2	78.2	83582	2	AC121741	AC128268 Rattus no
C 34	17.2	78.2	100499	9	AL138769	AC121741 Rattus no
C 35	17.2	78.2	102818	9	AC003042	AL138769 Human DNA
C 36	17.2	78.2	110000	2	LMFLCHR18_07	AC003042 Homo sapi
C 37	17.2	78.2	141877	10	AL669836	Continuation (8 of
C 38	17.2	78.2	161601	3	AC007594	AL669836 Mouse DNA
C 39	17.2	78.2	163386	9	AC015723	AC007594 Drosophill
C 40	17.2	78.2	171110	2	AC096216	AC015723 Homo sapi
C 41	17.2	78.2	171537	2	AC129925	AC096216 Rattus no
C 42	17.2	78.2	177564	2	AC119477	AC129925 Homo sapi
C 43	17.2	78.2	182381	2	AC096410	AC119477 Rattus no
C 44	17.2	78.2	197406	2	AC110594	AC096410 Rattus no
C 45	17.2	78.2	200158	2	AC116771	AC110594 Homo sapi
						AC116771 Mus muscu

ALIGNMENTS

RESULT	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE	AUTHORS	TITLE
1	AR171866	Sequence 7 from patent US 6297364.	AR171866	AR171866.1	GI:17910816	Unknown.	Unknown.	1 (bases 1 to 1946)	Chen, Y.-T., Gure, A., Teang, S., Stockert, E., Jager, E., Alexander, K. and Old, L.J.	Isolated nucleic acid molecule encoding cancer associated antigen, the antigen itself, and uses thereof
1	AR171866	Sequence 7 from patent US 6297364.	AR171866	AR171866.1	GI:17910816	Unknown.	Unknown.	1 (bases 1 to 1946)	Chen, Y.-T., Gure, A., Teang, S., Stockert, E., Jager, E., Alexander, K. and Old, L.J.	Isolated nucleic acid molecule encoding cancer associated antigen, the antigen itself, and uses thereof

Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Talamas, J.,
 Tesfaye, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H.,
 Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J.,
 Zembek, L., Zimmer, A. and Zody, M.

Direct Submission
 Submitted (06-AUG-2002) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA
 On Aug 6, 2002 this sequence version replaced gi:21426162.
 All repeats were identified using RepeatMasker:
 Smit, A.F.A. & Green, P. (1996-1997)
 http://ftp.genome.washington.edu/RM/RepeatMasker.html

----- Genome Center
 Center: Whitehead Institute/ MIT Center for Genome Research
 Center code: WIBR
 Web site: http://www-seq.wi.mit.edu
 Contact: sequence_submissions@genome.wi.mit.edu

----- Project Information
 Center project name: L23011
 Center clone name: 2244_F_11

 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 3 contigs. Gaps between the contigs
 * are represented as runs of N. The order of the pieces
 * is believed to be correct as given, however the sizes
 * of the gaps between them are based on estimates that have
 * provided by the submittor.
 * This sequence will be replaced
 * by the finished sequence as soon as it is available and
 * the accession number will be preserved.
 * 1 64310: contig of 64310 bp in length
 * 64311 64410: gap of 100 bp
 * 64411 86773: contig of 22363 bp in length
 * 86774 86873: gap of 100 bp
 * 86874 147767: contig of 60894 bp in length.

----- Location/Qualifiers
 1. 147767
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="17"
 /map="17"
 /clone="CTD-2244F11"
 /clone_lib="CITDI Human BAC"
 41792 a 34005 c 33361 g 38110 t 495 others

BASE COUNT 41792 a 34005 c 33361 g 38110 t

ORIGIN

Query Match 100.0%; Score 22; DB 2; Length 147767;
 Best Local Similarity 100.0%; Pred. No. 27;
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCCTTGGCGGCTGGCGGCTCTCAG 22
 Db 130408 TCCTTGGCGGCTGGCGGCTCTCAG 130387

RESULT 3
 AC091133/c 168613 bp DNA linear PRI 01-JUL-2002
 LOCUS Homo sapiens chromosome 17, clone RP11-501C14, complete sequence.
 DEFINITION AC091133
 ACCESSION AC091133.11 GI:21637524
 VERSION AC091133.11
 KEYWORDS HTG.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 168613)
 AUTHORS Birren, B., Nusbaum, C. and Lander, E.
 TITLE Homo sapiens chromosome 17, clone RP11-501C14
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 168613)
 AUTHORS Birren, B., Linton, L., Nusbaum, C., Lander, E., Allen, N., Anderson, S.,
 Barna, N., Bastien, V., Boguslavskiy, L., Boukhgalter, B., Brown, A.,
 Camarata, J., Chang, J., Campopiano, A., Chang, J., Choepel, Y., Colangelo, M.,
 Cooke, P., DeArrellano, K., Dewar, K., Diaz, J.S., Dodge, S.,
 Faroo, S., Ferreira, P., Fitzhugh, W., Gage, D., Galagan, J., Gardyna, S.,
 Ginde, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N.,
 Hones, B., Heaford, A., Horton, L., Hulme, W., Iliev, I., Johnson, R.,
 Jones, C., Katat, A., Karatas, A., Kells, C., LaRocque, K.,
 Lamarez, R., Landers, T., Lehoczy, J., Levine, R., Liu, G.,
 MacLean, C., Macdonald, P., Major, J., Marquis, N., Matthews, C.,
 McCarthy, M., McEwan, P., McKernan, K., McPheeters, R., Meldrum, J.,
 Meneus, L., Mihova, T., Mienga, V., Murphy, T., Naylor, J., Nguyen, C.,
 Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neil, D.,
 Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V.,
 Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P.,
 Roman, J., Rosetti, P., Roy, A., Santos, R., Schauer, S., Schupback, R.,
 Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N.,
 Straus, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J.,
 Topham, K., Travers, M., Travis, N., Trigilio, J., Vassiliev, H.,
 Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J., Young, G.,
 Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

Direct Submission
 Submitted (22-DEC-2001) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA
 3 (bases 1 to 147767)
 Birren, B., Nusbaum, C., Lander, E., Ali, A., Allen, N., Anderson, S.,
 Barna, N., Bastien, V., Bloom, T., Boguslavskiy, L., Boukhgalter, B.,
 Camarata, J., Chang, J., Chazaro, B., Choepel, Y., Collymore, A., S.,
 Cooke, P., DeArrellano, K., Dewar, K., Diaz, J.S., Dodge, S.,
 Faroo, S., Ferreira, P., Fitzhugh, W., Gage, D., Galagan, J.,
 Gardyna, S., Gord, S., Graham, L., Grand-Pierre, N., Hages, B.,
 Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Katat, A.,
 Karatas, A., Kells, C., Macdonald, P., Major, J., Matthews, C.,
 Liu, G., MacLean, C., Macdonald, P., Levine, R., Lindblad-Toh, K.,
 McCarthy, M., Meldrum, J., Meneus, L., Mihova, T., Mienga, V.,
 Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., Norman, C.H.,
 O'Connor, T., O'Donnell, P., O'Neil, D., Oliver, J., Peterson, K.,
 Phunkhang, P., Pierre, N., Raymond, C., Retta, R., Rise, C., Rogov, P.,
 Roman, J., Roy, A., Schauer, S., Schupback, R., Seaman, S., Severy, P.,

JOURNAL Patent: US 6297364-A 7 02-OCT-2001;
 Location/Qualifiers
 1. 1946
 /organism="unknown"

BASE COUNT 502 a 528 c 553 g 358 t 5 others

ORIGIN

Query Match 100.0%; Score 22; DB 6; Length 1946;
 Best Local Similarity 100.0%; Pred. No. 98;
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCCTTGGCGGCTGGCGGCTCTCAG 22
 Db 38 TCCTTGGCGGCTGGCGGCTCTCAG 59

RESULT 2
 AC105030/c 147767 bp DNA linear HTG 06-AUG-2002
 LOCUS Homo sapiens chromosome 17 clone CTD-2244F11 map 17, *** SEQUENCING
 DEFINITION IN PROGRESS ***, 3 ordered pieces.
 ACCESSION AC105030
 VERSION AC105030.4 GI:22123404
 KEYWORDS HTG; HTGS_PHASE2; HTGS_FULITOP; HTGS_ACTIVEFIN.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 147767)
 AUTHORS Birren, B., Nusbaum, C. and Lander, E.
 TITLE Homo sapiens chromosome 17, clone CTD-2244F11
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 147767)
 AUTHORS Birren, B., Linton, L., Nusbaum, C., Lander, E., Ali, A., Allen, N.,
 Anderson, S., Barna, N., Bastien, V., Boguslavskiy, L., Boukhgalter, B.,
 Brown, A., Camarata, J., Campopiano, A., Chang, J., Chazaro, B.,
 Choepel, Y., Colangelo, M., Collins, S., Collymore, A., Cooke, A.,
 Cooke, P., DeArrellano, K., Dewar, K., Diaz, J.S., Dodge, S., Faroo, S.,
 Ferreira, P., Fitzhugh, W., Gage, D., Galagan, J., Gardyna, S.,
 Ginde, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N.,
 Hones, B., Heaford, A., Horton, L., Hulme, W., Iliev, I., Johnson, R.,
 Jones, C., Katat, A., Karatas, A., Kells, C., LaRocque, K.,
 Lamarez, R., Landers, T., Lehoczy, J., Levine, R., Liu, G.,
 MacLean, C., Macdonald, P., Major, J., Marquis, N., Matthews, C.,
 McCarthy, M., McEwan, P., McKernan, K., McPheeters, R., Meldrum, J.,
 Meneus, L., Mihova, T., Mienga, V., Murphy, T., Naylor, J., Nguyen, C.,
 Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neil, D.,
 Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V.,
 Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P.,
 Roman, J., Rosetti, P., Roy, A., Santos, R., Schauer, S., Schupback, R.,
 Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N.,
 Straus, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J.,
 Topham, K., Travers, M., Travis, N., Trigilio, J., Vassiliev, H.,
 Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J., Young, G.,
 Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

Direct Submission
 Submitted (22-DEC-2001) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA
 3 (bases 1 to 147767)
 Birren, B., Nusbaum, C., Lander, E., Ali, A., Allen, N., Anderson, S.,
 Barna, N., Bastien, V., Bloom, T., Boguslavskiy, L., Boukhgalter, B.,
 Camarata, J., Chang, J., Chazaro, B., Choepel, Y., Collymore, A., S.,
 Cooke, P., DeArrellano, K., Dewar, K., Diaz, J.S., Dodge, S.,
 Faroo, S., Ferreira, P., Fitzhugh, W., Gage, D., Galagan, J.,
 Gardyna, S., Gord, S., Graham, L., Grand-Pierre, N., Hages, B.,
 Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Katat, A.,
 Karatas, A., Kells, C., Macdonald, P., Major, J., Matthews, C.,
 Liu, G., MacLean, C., Macdonald, P., Levine, R., Lindblad-Toh, K.,
 McCarthy, M., Meldrum, J., Meneus, L., Mihova, T., Mienga, V.,
 Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., Norman, C.H.,
 O'Connor, T., O'Donnell, P., O'Neil, D., Oliver, J., Peterson, K.,
 Phunkhang, P., Pierre, N., Raymond, C., Retta, R., Rise, C., Rogov, P.,
 Roman, J., Roy, A., Schauer, S., Schupback, R., Seaman, S., Severy, P.,

Collins, S., Collymore, A., Cooke, P., DeArellano, K., Dewar, K., Diaz, J.S., Dodge, S., Faro, S., Ferreira, P., FitzHugh, W., Gage, D., Galagan, J., Gardyna, S., Ginde, S., Goyette, M., Graham, L., Grand-Pierre, N., Hagos, B., Heaford, A., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Karatas, A., LaRocque, K., Lamazares, R., Landers, T., Lehocsky, J., Levine, R., Liu, G., MacLean, C., Macdonald, P., Marquis, N., Matthews, C., McCarthy, M., McEwan, P., McKernan, K., McPeeters, R., Meldrim, J., Meneus, L., Mihova, T., Mienga, V., Murphy, T., Naylor, J., Nguyen, C., Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V., Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P., Santos, R., Schauer, S., Schuback, R., Seaman, S., Severi, P., Sougnuez, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Strauss, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J., Travers, M., Travis, N., Trigilio, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

TITLE

JOURNAL

REFERENCE

AUTHORS

Submitted (01-APR-2001) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 168613)

Birren, B., Linton, L., Nusbaum, C., Lander, E., Ali, A., Allen, N., Anderson, S., Barna, N., Bastien, V., Bloom, T., Boguslavskiy, L., Boukhgalter, B., Brown, A., Camarata, J., Campopiano, A., Chang, J., Chazaro, B., Choepel, Y., Colangelo, M., Collins, S., Collymore, A., Cook, A., Cooke, P., DeArellano, K., Dewar, K., Diaz, J.S., Dodge, S., Faro, S., Ferreira, P., FitzGerald, M., FitzHugh, W., Gage, D., Galagan, J., Gardyna, S., Ginde, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N., Hagos, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kells, C., LaRocque, K., Lamazares, R., Landers, T., Lehocsky, J., Levine, R., Lindblad-Toh, K., Liu, G., MacLean, C., Macdonald, P., Major, J., Marquis, N., Matthews, C., McCarthy, M., McEwan, P., McKernan, K., Meldrim, J., Meneus, L., Mihova, T., Mienga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V., Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P., Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schuback, R., Seaman, S., Severi, P., Spencer, B., Stange-Thomann, N., Stojanovic, N., Strauss, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M., Travis, N., Trigilio, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

Direct Submission

TITLE

JOURNAL

REFERENCE

AUTHORS

Submitted (17-JUN-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
4 (bases 1 to 168613)

Birren, B., Nusbaum, C., Lander, E., Ali, A., Allen, N., Anderson, S., Barna, N., Bastien, V., Bloom, T., Boguslavskiy, L., Boukhgalter, B., Camarata, J., Chang, J., Chazaro, B., Choepel, Y., Collymore, A., Cook, A., Cooke, P., DeArellano, K., Dewar, K., Diaz, J.S., Dodge, S., Gardyna, S., Gord, S., Graham, L., Grand-Pierre, N., Hagos, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kells, C., Landers, T., Levine, R., Lindblad-Toh, K., Liu, G., MacLean, C., Macdonald, P., Major, J., Matthews, C., McCarthy, M., Meldrim, J., Meneus, L., Mihova, T., Mienga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Raymond, C., Retta, R., Rise, C., Rogov, P., Roman, J., Roy, A., Schauer, S., Schuback, R., Seaman, S., Severi, P., Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M., Travis, N., Trigilio, J., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

Direct Submission

TITLE

JOURNAL

COMMENT

Submitted (01-JUL-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
On Jul 1, 2002 this sequence version replaced gi:21431160.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: <http://www-seq.wi.mit.edu>
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: LI2028
Center clone name: 501_C_14

FEATURES

source	Location/Qualifiers
repeat_region	1..168613 /organism="Homo sapiens" /db_xref="taxon:9606" /chromosome="17" /map="17" /clones="RP11-501C14" /clone_lib="RPCI-11 Human Male BAC"
repeat_region	119..201 /rpt_family="(CA)n"
unsure	255..519 /rpt_family="AluJb" 456..460 /note="<30 quality SNGL region"
unsure	484..488 /note="<30 qual SNGL region"
repeat_region	1041..1349 /rpt_family="AluSx"
unsure	1271..1278 /note="<30 quality SNGL region"
unsure	1293..1296 /note="<30 quality SNGL region"
unsure	1394..1503 /note="single clone coverage"
unsure	1569..1573 /note="<30 quality SNGL region"
repeat_region	1592..1904 /rpt_family="AluSg"
unsure	1649..1653 /note="<30 quality SNGL region"
unsure	1672..1681 /note="<30 quality SNGL region"
unsure	1933..1937 /note="<30 quality SNGL region"
unsure	2002..2007 /note="<30 quality SNGL region"
repeat_region	2002..2007 /note="<30 quality SNGL region"
repeat_region	complement(3804..4099) /rpt_family="AluSp"
repeat_region	complement(4105..4156) /rpt_family="MER112"
repeat_region	4282..4316 /rpt_family="(T)n"
repeat_region	complement(4318..4415) /rpt_family="U6"
repeat_region	5302..5438 /rpt_family="MIR"
repeat_region	complement(5443..5828) /rpt_family="LIM4"
repeat_region	complement(6003..6309) /rpt_family="AluJb"
repeat_region	6349..6651 /rpt_family="AluSg"
repeat_region	6666..6941 /rpt_family="AluJo"
repeat_region	7125..7447 /rpt_family="AluSx"
repeat_region	7454..7628 /rpt_family="AluSc"
repeat_region	7656..7945 /rpt_family="AluJb"
repeat_region	8965..9277 /rpt_family="AluSx"
repeat_region	9365..9650 /rpt_family="AluSx"

[illegible]

```

/product="conserved hypothetical protein"
/protein_id="AAC65700.1"
/db_xref="GI:3323037"
/translation="MPCNTONLYQPGCKCTPMDEGRTVPQAHRAKEKKQDAHLAW
EVKREHACELRVFHVHOLESVSPKTVRFVTLTAPEWIVVPHVMEARQAFVWVR
QWEGCGTQVCTEFPGGVIDGHPPEAARRELFEETGRRASSLAHLGITHPNPAVLEN
RVHIFAECTPEVREFPOLDTDEFELRCVLPVHDVYRMGRAPFDHALMAQPSFFFCGR
IRFPNCNVRYVACSASI"
/complement(4007. .4954)
/gene="TP0732"
/gene="TP0732"
/complement(4007. .4954)
/gene="TP0732"
/notes="similar to SP:P54382 PID:1303916 GB:AL009126
percent identity: 47.14; identified by sequence
similarity; putative"
/codon_start=1
/transl_table=11
/product="methyltetrahydrofolate dehydrogenase (fold)"
/protein_id="AAC65701.1"
/db_xref="GI:3323038"
/translation="MDARLIDGQAHECTARLARTRVQALRAAVGTAPFLAALVGGDD
PASCYVAAKQALAHRLGTHRLPAHSHAQVLELIARLNEDARVHGILLOLPLP
AHLDAARVCRAVPEKVDGPHLNCGALFLAOPGFPCTPAGIWHLLRRQVFLAGA
RVVIVRSALVGRPLAVILLASPCDQATVTLCHSTRGLADICVQADILVLAALGKARFI
GAPFVRTGAVVDVGHVHPDAPRGRRLCGDVDFDAVAHKVQAITVPPGVGWPMTI
AMLLNLTCAAAYAGMIPFRALYADLDGSAAGDVPH"
/5050. .5709
/gene="TP0733"
/5050. .5709
/gene="TP0733"
/notes="hypothetical protein; identified by Glimmer;
putative"
/codon_start=1
/transl_table=11
/product="T. pallidum predicted coding region TP0733"
/protein_id="AAC65708.1"
/db_xref="GI:3323045"
/translation="MSRTFRAMQCVGALCALSPLLPAYSGEVREVPSPQVWVAY
EPIRPGDQLLXIGVAGQLYAGNGTNGSSGTNGKGLGGGHLGYEYFF
TKNFSLGQVGFYRTYYSFVITVNPYTFYAVGRWRIFLSGLVGLNIQSYLS
KXAPGLIAEAGAGLYYQYTPDWSIGGIVAYTQLGDIASSDPDKCRAVLATIDFGRVH
F"
/complement(5742. .6443)
/gene="TP0734"
/complement(5742. .6443)
/gene="TP0734"
/notes="similar to GP:1638807 percent identity: 57.51;
identified by sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="purine nucleoside phosphorylase (deod)"
/protein_id="AAC65702.1"
/db_xref="GI:3323039"
/translation="MSIHLAGKSDIAPRVLLPGDPLRARFVAERFFHNAYCYNEVRG
MLGFTGYKGVRSVQGTGMLPGSHSIVTTFODYGVQKARVGTAGGLQEEVALKS
VVCALGATSAIARVAFSGMHFAPTASMLTAVSVADEMTGTVGVGNVSSDFY
DESSGRLWARYGLVAVEMETAEYLTAAFRREALSILTIISHLVTVGAVTSAQERER
SFTOMIEALEAIQ"
/complement(6525. .8081)
/gene="TP0735"
/complement(6525. .8081)
/gene="TP0735"
/notes="similar to GP:2244618 percent identity: 48.79;
identified by sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="glutamate synthase (glta)"
/protein_id="AAC65703.1"
/db_xref="GI:3323040"
/translation="MTSVQNVVSDAKTCDVPVSEPEARTYRPHALQBEACQRLAQL
QSKELKMKDQIQLQRMMLPKKSLLMQEAALGFTQQALVESQRCINCKTKPCV
KGPVGVPIFEFTACVORGAFKEAVDILITLSLPAICRGVCPHCQLOCTVGMF
KDVSKAVSIGALERFVADMERQHGQITVPYCAPSTHKVAVIGSGPAGLAVASDTARA

```

```

CHSVTVFEALHKPGCVTYGIPERLPKRVVVVEIETLKKMGVTFRMFLVGRATLLE
QLFSQYGDVAVFIGTAGLPRFMNIEGSELGCVFAANDYLTRATLMKAYDTAHADTPV
YAASVWVVGNDVAVDSRTALRGAEQVHCLYKTRADMTACVBEIAQAEDEGVTF
HFLQCTTRILGDEEGNVAVVFRDCQIQIDAGERVFLPCGDAEACKLAADAVIVAGNG
SNPLMAKTRLSAVSERGTIVDEQRTSIPGVWAGGDIVLGAATVIRMGCGRRAA
SINPLAQAQA"
/complement(8084. .8926)
/gene="TP0736"
/complement(8084. .8926)
/gene="TP0736"
/notes="similar to GB:L77117 PID:1592091 percent identity:
36.32; identified by sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="hydrogenase, gamma chain (hydG)"
/protein_id="AAC65704.1"
/db_xref="GI:3323041"
/translation="MFSILEKKQFSPQVFLVKVHAPDIKNRAAGOFVLVOLDDEYAE
RPLTIADAHADSEGWIALVIQVGTANTMLCEKEVGDSSVVLGPNPLIENGVTV
ACVAGGVGAAPLYPIAQAHKRAHNVIVILGARNRDLIIFEEEMRALDELIVITDDG
SYRKGVLTEPLRELCKARASCPQEVVAIGPIMMKFCAETTRFPGIRTVVSLNTIMID
GTGKGGCGCRVNVGSEIKFVCVDGPEFDGHQVDFESMMRRMRAPQAREAHVHCCMKH
HAGR"
/9160. .10470
/gene="TP0737"
/9160. .10470
/gene="TP0737"
/notes="similar to GP:2293308 percent identity: 22.83;
identified by sequence similarity; putative"
/codon_start=1

Query Match      85.5%; Score 18.8; DB 1; Length 12110;
Best Local Similarity 90.9%; Pred. No. 9.9e+02;
Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 TCCTTGGCGCGCTGCGGCTCAG 22
    ||| ||||| ||||| |||||
Db 3729 TCCTTGGCGCGCTGCGGCTCAG 3708

RESULT 5
AC118983/c
LOCUS      AC118983      175585 bp      DNA      linear      HTG 23-APR-2002
DEFINITION Rattus norvegicus clone RP31-438E3 strain Brown Norway, WORKING
            DRAFT SEQUENCE, 19 unordered pieces.
AC118983
AC118983.1 GI:20270129
VERSION     HTG; HTGS_PHASE1; HTGS_DRAFT.
KEYWORDS   Rattus norvegicus.
SOURCE     Rattus norvegicus
ORGANISM   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE  1 (bases 1 to 175585)
            Akhter,N., Antonellis,A., Ayele,K., Beckstrom-Sternberg,S.M.,
            Benjamin,B., Blakesley,R.W., Bouffard,G.G., Breen,K., Brinkley,C.,
            Brooks,S., Dietrich,N.L., Granite,S., Guan,X., Gupta,J.,
            Haghighi,P., Hansen,N., Ho,S.-L., Idol,J.R., Karlins,E., Laric,P.,
            Lee-Lin,S.-Q., Legaapi,R., Maduro,Q.L., Maduro,V.B.,
            Margulies,E.H., Mastello,C., Maskeri,B., Mastrian,S.D.,
            McCloskey,J.C., McDowell,J., Pagnirigan,C., Pearson,R.,
            Porcnoy,M.E., Prasad,A., Schueler,M.G., Stantropop,S., Thomas,J.W.,
            Thomas,P.J., Touchman,J.W., Taurgeon,C., Vogt,J.L., Walker,M.A.,
            Wetherby,K.D., Wiggins,L., Young,A., Zhang,L.-H. and Green,E.D.,
            NISC Comparative Sequencing Initiative
            Unpublished
            2. (bases 1 to 175585)
            Green,E.D.
            Direct Submission
            Submitted (23-APR-2002) NIH Intramural Sequencing Center, 8717
            Grovemont Circle, Gaithersburg, MD 20877, USA
            ----- Genome Center
            Center: NIH Intramural Sequencing Center

```

```

Center code: NISC
Web site: http://www.nisc.nih.gov
Contact: nisc.zoo@hghri.nih.gov
----- Project Information -----
Center project name: dcj
Center clone name: 438E03
----- Summary Statistics -----
Sequencing vector: plasmid; n/a; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 168197 bases at least Q40
Consensus quality: 169738 bases at least Q30
Consensus quality: 170660 bases at least Q20
Insert size: 200000; agarose-fp
Insert size: 173785; sum-of-contents
Quality coverage: 4.54x in Q20 bases; agarose-fp
Quality coverage: 5.23x in Q20 bases; sum-of-contents

```

```

* NOTE: This is a 'working draft' sequence. It currently
* consists of 19 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of 'N', but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

```

* be preserved.	2278:	contig	of 2278 bp	in length
*	1			
*	2279	2378:	gap of unknown length	
*	2379	4780:	contig of 2402 bp	in length
*	4781	4880:	gap of unknown length	
*	4881	8217:	contig of 3337 bp	in length
*	8218	8317:	gap of unknown length	
*	8318	11821:	contig of 3504 bp	in length
*	11822	11921:	gap of unknown length	
*	11922	16465:	contig of 4544 bp	in length
*	16466	16565:	gap of unknown length	
*	16566	22553:	contig of 5698 bp	in length
*	22554	22553:	gap of unknown length	
*	22554	28019:	contig of 5656 bp	in length
*	28020	28119:	gap of unknown length	
*	28120	33011:	contig of 4892 bp	in length
*	33012	33111:	gap of unknown length	
*	33112	36974:	contig of 3863 bp	in length
*	36975	37074:	gap of unknown length	
*	37075	43073:	contig of 5999 bp	in length
*	43074	43173:	gap of unknown length	
*	43174	50920:	contig of 7747 bp	in length
*	50921	51020:	gap of unknown length	
*	51021	61409:	contig of 10389 bp	in length
*	61410	61509:	gap of unknown length	
*	61510	71366:	contig of 9857 bp	in length
*	71367	71466:	gap of unknown length	
*	71467	83535:	contig of 12069 bp	in length
*	83536	83635:	gap of unknown length	
*	83636	99626:	contig of 15991 bp	in length
*	99627	99736:	gap of unknown length	
*	99727	113517:	contig of 13791 bp	in length
*	113518	113617:	gap of unknown length	
*	113618	123546:	contig of 9929 bp	in length
*	123547	123646:	gap of unknown length	
*	123647	152337:	contig of 28691 bp	in length
*	152338	152437:	gap of unknown length	
*	152438	175585:	contig of 23148 bp	in length.

FEATURES

```
1. .175585
/organism="Rattus norvegicus"
/strain="Brown Norway"
/db_xref="taxon:10116"
/_clone="RP31-438E3"
/_clone_lib="RP31"
```

```
misc_feature
1. .2278
/note="assembly_fragment"
2379. .4780
misc_feature
```

```

misc_feature      /note="assembly_fragment"
4881. .8217
/note="assembly_fragment"
clone_end:SP6
vector_side:right"
8318. .11821
/note="assembly_fragment"
11922. .16465
/note="assembly_fragment"
16566. .22253
/note="assembly_fragment"
22344. .28019
/note="assembly_fragment"
clone_end:T7
vector_side:right"
28120. .33011
/note="assembly_fragment"
33112. .36974
/note="assembly_fragment"
37075. .43073
/note="assembly_fragment"
43174. .50920
/note="assembly_fragment"
51021. .61409
/note="assembly_fragment"
61510. .71366
/note="assembly_fragment"
71467. .83535
/note="assembly_fragment"
83636. .99626
/note="assembly_fragment"
99721. .113517
/note="assembly_fragment"
113618. .123546
/note="assembly_fragment"
123647. .152337
/note="assembly_fragment"
152438. .175585
/note="assembly_fragment"
45269 a 40110 c 40996 g 47399
BASE COUNT
ORIGIN
```

Query Match	85.5%	Score 18.8;	DB 2;	Length 175585;
Best Local Similarity	90.9%	Pred. No. 4.4e+02;		
Matches	20;	Conservative	0;	Mismatches 2;
			Indels	0;
			Gaps	0;

QY 1 TCCTTGGCGGTGCGGCTCAG 22
pb 75043 TCCTTGGCGGTGCGGCCCCAG 75022

RESULT 6				PRI	31-JAN-2002
AF463492/c				linear	
LOCUS	AF463492	2033 bp	mRNA		
DEFINITION	Homo sapiens inhibitory FAS domain protein (IPAS)				complete
	cds.				
ACCESSION	AF463492				
VERSION	AF463492.1	GI:18448894			
KEYWORDS	.				
SOURCE	Homo sapiens.				

REFERENCE	Mammalia, Rethna, and Rethna. 2023. Cloning and characterization of human inhibitory PAS domain protein 1 (bases 1 to 2023).
AUTHORS	Cheng, J.Q.
TITLE	Cloning and characterization of human inhibitory PAS domain protein
JOURNAL	Unpublished
REFERENCE	2. (bases 1 to 2023)
AUTHORS	Cheng, J.Q.
TITLE	Direct Submission
JOURNAL	Submitted (27-DEC-2001) Pathology, University of South Florida,
FEATURES	12901 Bruce B. Downs Blvd., MDC11, Tampa, FL 33612, USA

ES
....., Institute of Medical Science,
University of Tokyo.
Location/Qualifiers
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/ILNL at: <http://image.llnl.gov>
Series: IRAK Plate: 32 Row: j Column: 19

Mon Apr 21 10:29:54 2003

This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 11968004.

FEATURES

source

1. .2477
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="LocusID:64344"
/db_xref="taxon:9606"
/clone="MGC:26043 IMAGE:4811824"
/tissue type="Brain, hippocampus"
/clone lib="NIH MGC_95"
/lab host="DH10B"
/note="Vector: pBluescript"

CDS

23. .1114
/codon_start=1
/product="hypothetical protein FLJ14033 similar to hypoxia inducible factor 3, alpha subunit"
/protein id="AAH26308.1"
/db_xref="GI:20072846"
/translation="MALGLQARSTTELIRKEKSRDAARRSQETEVLYQLAHTLPFA
RGVSHLDKASIMBLTISYLRMHLCAAGNNOVGAGEPLDACYLKALGFYVNLTA
EGDMAYLSNWSKHLGSLQLGHSIPDFTHPCDQELQDALTPQOTLSRRKVEART
ERCFSLRWKSLTSGRTLNKAKATKVLNCSGHWRAVKPPAOTSPPGSDPSPPLQC
LVLI CEAI PHPGSLEPPGLRGAFLSRHSLMKFTYCDRIAEGVAGSPDLIGCSAYE
YTHALSDAVKSIHTCMYPI SPGAKPAATWPPADTRTPQULPQDALPHLNTSSIL
PKPQCTVSFLAPSPVPRSPFPHPPMP"

BASE COUNT

577 a 659 c 684 g 557 t

ORIGIN

Query Match 80.9%; Score 17.8; DB 9; Length 2477;
Best Local Similarity 90.5%; Pred. No. 3.9e+03;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY

2 CCTTGGCGCTGGCGCTCAG 22

Db

49 CCTTGGCGCTGGCGCCCCAG 29

RESULT 9

AC007193/c 146180 bp DNA linear PRI 17-JUN-1999
LOCUS Homo sapiens chromosome 19, BAC 82621 (CIT-B-139a18), complete
DEFINITION sequence.

ACCESSION

AC007193

VERSION

AC007193.1 GI:4558635

KEYWORDS

HTG.

SOURCE

Homo sapiens.

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 146180)
Lamerdin,J.E., McCready,P.M., Skowronski,E., Viswanathan,V.,
Burkhardt-Schultz,K.J., Gordon,L., Dias,J., Ramirez,M., Terry,A.,
Stilwagen,S., Phan,H., Velasco,N., Do,L., Regala,W., Terry,A.,
Garnes,J., Danganan,L., Erler,A., Christensen,M., Georgescu,A.,
Avila,J., Liu,S., Attix,C., Andrade,T., Trankheim,M.,
Amico-Keller,G., Coefield,J., Duarte,S., Lucas,S., Bruce,R.,
Thomas,P., Quan,G., Kronmiller,B., Arellano,A., Sanders,C., Ow,D.,
Nolan,M., Trong,S., Kobayashi,A., Olsen,A.S. and Carrano,A.V.
Sequence analysis of a 1.9 Mb region in 19q13.2 between APOE and
D19S412

TITLE

Unpublished

JOURNAL

2 (bases 1 to 146180)

REFERENCE

Lamerdin,J.E.

AUTHORS

Submitted (02-APR-1999)

TITLE

National Laboratory, 7000 East Ave., Livermore, CA 94551, USA

JOURNAL

Map and sequence oriented from q centromere to telomere, BAC 82621

COMMENT

(CIT-B-139a18) overlaps BAC 264576 (CIT-B-297n14) to the left from

bases 1 to 27,487 of this accession, and separated from BAC 89581

(CIT-B-158d10) on the right by a gap of approximately 17 to 23 kb.

Additional chromosome 19 map and sequence information may be

obtained at: <http://www.bio.lnl.gov/bbrp/genome/genome.html>.

FEATURES

Location/Qualifiers

source

1. .146180
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="19"
/map="19q13.2 between APOE and D19S412"
/clone="BC82621"
/cell line="fibroblast"
/note="LLNL clone name: BC82621 BAC library obtained from
Research Genetics."
1. .101
/rpt family="L1"
repeat_region
511. .677
/rpt family="L1"
repeat_region
1018. .1158
/rpt family="L1"
repeat_region
1159. .1453
/rpt family="AluSx"
repeat_region
1459. .1529
/rpt family="L1"
repeat_region
complement (1857. .1974)
/rpt family="FLAM_A"
2115. .2375
/rpt family="L1MB4"
repeat_region
2478. .2549
/rpt family="LINE2"
repeat_region
complement (2629. .2921)
/rpt family="AluJo"
3355. .3655
/rpt family="AluSx"
repeat_region
3673. .3939
/rpt family="L1"
repeat_region
complement (4005. .4264)
/rpt family="AluSx"
repeat_region
complement (4269. .4568)
/rpt family="AluY"
5064. .5122
/note="predicted exon, program: grail2exons_human_1.3,
frame: 2, quality: excellent, score: 81.000"
complement (5169. .5468)
/rpt family="AluY"
5626. .5886
/note="predicted exon, program: grail2exons_human_1.3,
frame: 0, quality: good, score: 52.000"
7398. .7444
/rpt family="Alu"
repeat_region
7480. .7595
/rpt family="TA)n"
repeat_region
7596. .7812
/rpt family="AluJb"
repeat_region
7900. .8191
/rpt family="AluSx"
repeat_region
complement (8575. .8597)
/rpt family="Ar_rich"
repeat_region
complement (8598. .8878)
/rpt family="AluSx"
repeat_region
9477. .9766
/rpt family="AluSx"
repeat_region
9997. .10258
/rpt family="AluSx"
repeat_region
10273. .10347
/rpt family="FLAM"
repeat_region
10691. .10818
/rpt family="AluJb"
repeat_region
complement (10999. .11283)
/rpt family="AluJo"
repeat_region
complement (11302. .11517)
/rpt family="L1MC/D"
repeat_region
complement (11536. .11661)
/rpt family="FLAM_C"
repeat_region
complement (11667. .11798)
/rpt family="L1MC/D"
repeat_region
complement (11833. .12138)

```

repeat_region      /rpt family="AluSg"
repeat_region      complement(12158. .12478)
repeat_region      /rpt family="L1MC/D"
repeat_region      12479. .12642
repeat_region      /rpt family="FRAM"
repeat_region      complement(12655. .12795)
repeat_region      /rpt family="L1MC/D"
repeat_region      complement(12806. .13106)
repeat_region      /rpt family="AluSx"
repeat_region      complement(13356. .13481)
repeat_region      /rpt family="GAAA)n"
repeat_region      complement(13482. .13694)
repeat_region      /rpt family="AluJb"
repeat_region      complement(13709. .14037)
repeat_region      /rpt family="L1"
repeat_region      complement(14009. .14105)
repeat_region      /rpt family="L1MC/D"
repeat_region      complement(14115. .14415)
repeat_region      /rpt family="AluY"
repeat_region      complement(14417. .14849)
repeat_region      /rpt family="L1MC/D"
repeat_region      complement(14859. .15153)
repeat_region      /rpt family="AluSx"
repeat_region      15226. .15492
repeat_region      /rpt family="AluJo"
repeat_region      complement(16257. .16340)
repeat_region      /note="predicted exon, program: grill2exons_human_1.3,
repeat_region      frame: 0, quality: good, score: 63.000"
repeat_region      16362. .16666
repeat_region      /rpt family="AluSg"
repeat_region      complement(16691. .17060)
repeat_region      /rpt family="THE1C"
repeat_region      17482. .17781
repeat_region      /rpt family="AluJo"
repeat_region      17782. .17870
repeat_region      /rpt family="n"
repeat_region      17968. .18273
repeat_region      /rpt family="AluSg"
repeat_region      complement(18401. .18817)
repeat_region      /rpt family="LTR7"
repeat_region      complement(18818. .19193)
repeat_region      /rpt family="HERVH"
repeat_region      complement(19194. .19412)
repeat_region      /rpt family="HERVH"
repeat_region      complement(19414. .20028)
repeat_region      /rpt family="HERVH"
repeat_region      complement(20031. .23647)
repeat_region      /rpt family="HERVH"
repeat_region      complement(23651. .24066)
repeat_region      /rpt family="LTR7"
repeat_region      complement(24403. .24714)
repeat_region      /rpt family="MER74"
repeat_region      24789. .25080
repeat_region      /rpt family="AluSx"
repeat_region      25161. .25207
repeat_region      /rpt family="L1MD3"
repeat_region      25224. .25509
repeat_region      /rpt family="AluJb"
repeat_region      25605. .25901
repeat_region      /rpt family="AluSp"
repeat_region      25902. .26069
repeat_region      /rpt family="AluSg/x"
repeat_region      complement(26224. .26256)
repeat_region      /rpt family="TAA)n"
repeat_region      complement(26257. .26545)
repeat_region      /rpt family="AluJo"
repeat_region      26573. .26876
repeat_region      /rpt family="AluJo"
repeat_region      complement(27125. .27422)
repeat_region      /rpt family="AluSx"
repeat_region      27866. .28173
repeat_region      /rpt family="AluSx"
repeat_region      28181. .28279

```

```

repeat_region      /rpt family="AluJ/FLAM"
repeat_region      complement(28359. .28522)
repeat_region      /rpt family="AluSg"
repeat_region      complement(28526. .28705)
repeat_region      /rpt family="AluSc"
repeat_region      complement(28733. .28872)
repeat_region      28874. .29081
repeat_region      /rpt family="AluJo"
repeat_region      29463. .29583
repeat_region      /note="predicted exon, program: grill2exons_human_1.3,
repeat_region      frame: 2, quality: good, score: 71.000"
repeat_region      complement(29585. .29888)
repeat_region      /rpt family="AluSg"
repeat_region      30336. .30634
repeat_region      /rpt family="AluSg"
repeat_region      30646. .30779
repeat_region      /rpt family="AluJb"
repeat_region      complement(30830. .31040)
repeat_region      /rpt family="MIR"

Query Match      80.9%; Score 17.8; DB 9; Length 146180;
Best Local Similarity 90.5%; Pred.No. 1.1e+03;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 CTTTGGCGCTGCGGCTCAG 22
Db 42317 CTTTGGCGCTGCGGCTCAG 42297

```

RESULT 10

AC099392

LOCUS

DEFINITION

Rattus norvegicus clone CH230-160L8, *** SEQUENCING IN PROGRESS

AC099392

AC099392.6 GI:21744099

VERSION

HTG: HTGS PHASE1.

KEYWORDS

Norway rat.

ORGANISM

Rattus norvegicus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;

Rattus.

REFERENCE

1 (bases 1 to 210868)

Muzny, D.M., Adams, C., Adio-Oduola, B., Ali-Osman, F.R., Allen, C.,

Alsbrooks, S.L., Amaratunge, H.C., Are, J.R., Ayale, M., Banks, T.,

Barbata, J., Benton, J., Bimage, K., Blankenburg, K., Bonnin, D.,

Bouck, J., Bowie, S., Brivea, M., Brown, E., Brown, M., Bryant, N.P.,

Buhay, C., Burch, P., Burkett, C., Burrell, K.L., Byrd, N.C.,

Carron, T.F., Carter, M., Cavazos, S.R., Chacko, J., Chavez, D.,

Chen, G., Chen, R., Chen, Z., Chowdhry, I., Christopoulos, C.,

Cleveland, C.D., Cox, C., Coyle, M.D., Dathorne, S.R., David, R.,

Davila, M.L., Davis, C., Davy-Carroll, L., Dederich, D.A.,

Douthwaite, K.J., Draper, H., Dugan-Rocha, S., Durbin, K.J.,

Earnhart, C., Edgar, D., Edwards, C.C., Elhaj, C., Escotto, M.,

Falls, T., Ferraguto, D., Flagg, N., Ford, J., Foster, P., Frantz, P.,

Gabis, A., Gao, J., Garcia, A., Garner, T., Garza, N., Gill, R.,

Gorrell, J.H., Guevara, W., Gunaratne, P., Hale, S., Hamilton, K.,

Harris, C., Harris, K., Hart, M., Havlak, P., Hawes, A., Hernandez, J.,

Hernandez, O., Hodgson, A., Hughes, M., Holloway, C., Hollins, B.,

Honsi, F., Howard, S., Huber, J., Hulyk, S., Hume, J., Jackson, L.E.,

Jacobson, B., Jia, Y., Johnson, R., Jolivet, S., Joudah, S.,

Karlsson, E., Kelly, S., Khan, U., King, L., Korvah, J., Kovar, C.,

Kratovic, J., Kureshi, A., Landry, N., Leal, B., Lewis, L.C., Lewis, L.,

Li, J., Li, Z., Lichtarge, O., Lieu, C., Liu, J., Liu, W., Loulseged, H.,

Lozano, R.J., Lu, X., Lucier, A., Lucier, R., Luna, R., Ma, J.,

Maneshwari, M., Mapua, P., Martin, R., Martindale, A., Martinez, E.,

Massey, E., Mawhiney, E., McLeod, M.P., Meador, M., Mei, G., Metzger, M.,

Miner, G., Miner, Z., Mitchell, T., Mohabbat, K., Morgan, M., Morris, N.,

Moser, M., Neal, D., Newton, J., Newton, N., Nguyen, A., Nguyen, N.,

Nguyen, N., Nickerson, E., Nwokenkwo, S., Ogih, M., Okwuonu, G.,

Oragunye, N., Oviedo, R., Pace, A., Payton, B., Peery, J., Perez, L.,

Mon Apr 21 10:29:54 2003

Peters, L., Pickens, R., Primus, E., Pu, L.L., Quiles, M., Ren, Y.,
Rives, M., Rojas, A., Rojokan, I., Rolfe, M., Ruiz, S., Savary, G.,
Scherer, S., Scott, G., Shen, H., Shoohrtari, N., Sisson, I.,
Sodergren, E., Sonaik, T., Sparks, A., Stanley, H., Stone, H.,
Sutton, A., Svatek, A., Tabor, P., Tamerisa, A., Tamerisa, K., Tang, H.,
Tansey, J., Taylor, C., Taylor, T., Telford, B., Thomas, N., Thomas, S.,
Usman, K., Vasquez, L., Vera, V., Villalon, D., Vinson, R., Wang, O.,
Wang, S., Ward-Moore, S., Warren, R., Washington, C., Watlington, S.,
Williams, G., Williamson, A., Wlezyk, R., Woodson, S., Worley, K.,
Wu, C., Wu, Y.F., Zhou, J., Zorrilla, S., Nelson, D.,
Weinstock, G. and Gibbs, R.
Direct Submission
Unpublished
2 (bases 1 to 210868)
Worley, K.C.
Direct Submission
Submitted (10-NOV-2001) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 210868)
Worley, K.C.
Direct Submission
Submitted (17-JUN-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Jul 13, 2002 this sequence version replaced gi:18700960.
----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GKZE
Center clone name: CH230-160L8
----- Summary Statistics
Sequencing vector: Plasmid
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 147538 bases at least Q40
Consensus quality: 152247 bases at least Q30
Consensus quality: 155330 bases at least Q20

* NOTE: Estimated insert size may differ from sequence length
(see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 69 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1 1434: contig of 1434 bp in length
* 1435: gap of unknown length
* 1535: contig of 1124 bp in length
* 2659: gap of unknown length
* 4175: contig of 1416 bp in length
* 4275: gap of unknown length
* 5457: contig of 1182 bp in length
* 5557: gap of unknown length
* 6846: contig of 1289 bp in length
* 6946: gap of unknown length
* 8314: contig of 1368 bp in length
* 8414: gap of unknown length
* 9594: contig of 1080 bp in length
* 11067: gap of unknown length
* 12116: contig of 1473 bp in length
* 12215: gap of unknown length
* 12316: contig of 1049 bp in length
* 13613: gap of unknown length
* 13712: contig of 1297 bp in length
* 13713: gap of unknown length
* 15322: contig of 1610 bp in length
*
* 15422: gap of unknown length
* 16673: contig of 1251 bp in length
* 16773: gap of unknown length
* 18285: contig of 1512 bp in length
* 18385: gap of unknown length
* 19335: contig of 1250 bp in length
* 19336: gap of unknown length
* 21369: contig of 1634 bp in length
* 21469: gap of unknown length
* 22777: contig of 1308 bp in length
* 22778: gap of unknown length
* 24514: contig of 1637 bp in length
* 24515: gap of unknown length
* 26098: contig of 1484 bp in length
* 26199: gap of unknown length
* 28120: contig of 1922 bp in length
* 28121: gap of unknown length
* 28221: contig of 1262 bp in length
* 29482: gap of unknown length
* 29582: contig of 1383 bp in length
* 30966: gap of unknown length
* 30966: contig of 1374 bp in length
* 32439: contig of 2102 bp in length
* 32440: gap of unknown length
* 32540: contig of 1198 bp in length
* 34741: gap of unknown length
* 35339: contig of 1241 bp in length
* 36039: gap of unknown length
* 36040: contig of 1241 bp in length
* 37281: gap of unknown length
* 37281: contig of 2628 bp in length
* 37381: gap of unknown length
* 40108: contig of 2518 bp in length
* 42626: gap of unknown length
* 42726: contig of 2081 bp in length
* 44807: gap of unknown length
* 44907: contig of 1380 bp in length
* 46387: gap of unknown length
* 48246: gap of unknown length
* 48346: contig of 1859 bp in length
* 50359: gap of unknown length
* 50459: contig of 2013 bp in length
* 51749: gap of unknown length
* 51749: contig of 1290 bp in length
* 51849: gap of unknown length
* 51917: contig of 2068 bp in length
* 54017: gap of unknown length
* 56173: contig of 2156 bp in length
* 56273: gap of unknown length
* 58609: contig of 2336 bp in length
* 58709: gap of unknown length
* 61307: contig of 2598 bp in length
* 61407: gap of unknown length
* 62553: contig of 1146 bp in length
* 62553: gap of unknown length
* 62554: contig of 2933 bp in length
* 65586: gap of unknown length
* 65587: contig of 2037 bp in length
* 67723: gap of unknown length
* 67823: contig of 1883 bp in length
* 69706: gap of unknown length
* 69806: contig of 2206 bp in length
* 72012: gap of unknown length
* 72112: contig of 1611 bp in length
* 73723: gap of unknown length
* 73824: contig of 2299 bp in length
* 76122: gap of unknown length
* 76222: contig of 2357 bp in length
* 78579: gap of unknown length
* 78579: contig of 4060 bp in length
* 82739: gap of unknown length
* 82839: contig of 3541 bp in length
* 86380: gap of unknown length
* 86480: contig of 2953 bp in length
* 89433: gap of unknown length
* 89533: gap of unknown length

```

* 89534 92623: contig of 3090 bp in length
* 92624 92723: gap of unknown length
* 92724 96987: contig of 4264 bp in length
* 96988 97087: gap of unknown length
* 97088 98133: contig of 2046 bp in length
* 98134 99233: gap of unknown length
* 99234 103909: contig of 4676 bp in length
* 103910 104009: gap of unknown length
* 104010 108102: contig of 4093 bp in length
* 108103 108202: gap of unknown length
* 108203 110791: contig of 2589 bp in length
* 110792 110891: gap of unknown length
* 110892 113718: contig of 2827 bp in length

Query Match      80.9%; Score 17.8; DB 2; Length 210868;
Best Local Similarity 90.5%; Pred. No. 1e+03; Indels 0; Gaps 0;
Matches 19; Conservative 0; Mismatches 2;

QY 1 TCCTTGGCGCTGCGGCTCA 21
Db 27677 TCCTTACGTGCTGGGCTCA 27697

RESULT 11
R1CG148A/c
LOCUS
DEFINITION
Oryza sativa (japonica cultivar-group) 352 bp DNA linear STS 29-MAY-2002
RFLP marker, clone G148, sequence tagged site.
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzaceae; Oryza.
REFERENCE
1 (bases 1 to 352)
AUTHORS
Minobe,Y.
TITLE
Nucleotide sequence of Sequence Tagged Site from rice RFLP marker
probe
JOURNAL
Unpublished
COMMENT
PROJECT = 'RGP'.
FEATURES
Location/Qualifiers
source
1..352
/organism="Oryza sativa (japonica cultivar-group)"
/cultivar="Nipponbare"
/db_xref="taxon:39947"
/chromosome="12"

BASE COUNT      63 a 63 c 109 g 117 t
ORIGIN

Query Match      79.1%; Score 17.4; DB 11; Length 352;
Best Local Similarity 94.7%; Pred. No. 1e+04; Indels 0; Gaps 0;
Matches 18; Conservative 0; Mismatches 1;

QY 1 TCCTTGGCGCTGCGGCTT 19
Db 315 TCCTTGGCGCTGCTGCTT 297

RESULT 12
AB061849
LOCUS
DEFINITION
Homo sapiens RPL22 gene for ribosomal protein L22, exons 1, 2 and
partial cds.
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Homo sapiens DNA, clone lib:Keio BAC library.
Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1

```

Yoshihama,M., Uechi,T., Asakawa,S., Kawasaki,K., Kato,S., Higa,S.,
Maeda,N., Minoshima,S., Tanaka,T., Shimizu,N. and Kenmochi,N.,
The human ribosomal protein genes: sequencing and comparative
analysis of 73 genes
Genome Res. 12 (3), 379-390 (2002)

JOURNAL
MEDLINE
REFERENCE
21864036
2 (bases 1 to 7419)
Kenmochi,N. and Shimizu,N.
Direct Submission
Submitted (17-MAY-2001) Naoya Kenmochi, Miyazaki Medical College,
Central Research Laboratories, 5200 Kihara, Kiyosake, Miyazaki
889-1692 Japan (E-mail:kenmochi@post.miyazaki-med.ac.jp.
Tel:81-985-85-9665, Fax:81-985-85-1514)

FEATURES
Location/Qualifiers
source
1..7419
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="1"
/map="1p36.3-p36.2"
/clone_lib="Keio BAC library"
85..187
/rpt_family="MERSA"
complement(229..298)
/rpt_family="L2"
complement(302..386)
/rpt_family="L2"
816..7419
/gene="RPL22"
816..850
/gene="RPL22"
/number=1
join(839..850,2657..>2761)
/gene="RPL22"
/codon_start=1
/product="ribosomal protein L22"
/protein_id="BAB79487.1"
/db_xref="GI:17932992"
/translation="MAPVKLVVGGKKKKQVLKFTLDCTHPVEDGIMDAANP"
851..2656
/gene="RPL22"
/number=1
2657..2761
/gene="RPL22"
/number=2
2762..>7419
/gene="RPL22"
/number=2
3084..3383
/rpt_family="AluSg"
complement(4321..4632)
/rpt_family="AluSp"
complement(4634..4945)
/rpt_family="AluSx"
complement(4949..5249)
/rpt_family="AluY"
complement(5255..5519)
/rpt_family="AluY"
complement(5520..5808)
/rpt_family="AluJo"
complement(5932..6028)
/rpt_family="MIR3"
complement(6187..6471)
/rpt_family="AluJb"
6663..6791
/rpt_family="AluJb"
6792..7089
/rpt_family="AluSp"
7090..7267
/rpt_family="AluJb"
1705 a 1683 c 1882 g 2149 t

Query Match 79.1%; Score 17.4; DB 9; Length 7419;

```

/clones="OJ1410_G05"
/clone lib="Monsanto"
BASE COUNT 23697 a 17081 c 17645 g 24222 t 407 others
ORIGIN
Query Match 79.1%; Score 17.4; DB 2; Length 83052;
Best Local Similarity 94.7%; Pred. No. 1.9e+03;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TCCTTGGCGCTGGCGCCT 19
|||||
Db 66870 TCCTTGGCGCTGGCTGCCT 66888

RESULT 14
AL589792 132592 bp DNA linear HTG 10-JUL-2001
LOCUS Homo sapiens chromosome 1 clone RP5-889P23 map p36.21-36.33, ***
DEFINITION SEQUENCING IN PROGRESS ***, 4 unordered pieces.
ACCESSION AL589792
VERSION AL589792.1 GI:13277194
KEYWORDS HTG; HTGS_PHASE1; HTGS_CANCELLED.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 132592)
AUTHORS Lloyd, C.
TITLE Direct Submission
JOURNAL Submitted (10-JUL-2001) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk
requests: clonerequest@sanger.ac.uk
COMMENT ----- Genome Center
Center: Sanger Centre
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: humquery@sanger.ac.uk
----- Project Information
Center project name: dj889P23
----- Summary Statistics
Assembly program: XGAP4; version 4.5
Sequencing vector: M13; M77815; 12% of reads
Sequencing vector: plasmid; L08752; 87% of reads
Chemistry: Dye-terminator ABI; 14% of reads
Chemistry: Dye-terminator Big Dye; 60% of reads
Chemistry: Dye-terminator-amersham; 24% of reads
Consensus quality: 131458 bases at least Q40
Consensus quality: 131709 bases at least Q30
Consensus quality: 131965 bases at least Q20
Insert size: 132292; sum-of-contigs
Insert size: 132129; 14.2% error; agarose-fp
Quality coverage: 11.37x in Q20 bases; sum-of-contigs Quality
coverage: 11.69x in Q20 bases; agarose-fp
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 4 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 3331: contig of 3331 bp in length
* 3332 3431: gap of 100 bp
* 3432 90998: contig of 87567 bp in length
* 90999 91098: gap of 100 bp
* 91099 130106: contig of 39008 bp in length
* 130107 130206: gap of 100 bp
* 130207 132592: contig of 2386 bp in length.
* Location/Qualifiers
FEATURES
source 1..132592
/organism="Homo sapiens"

```

```

Best Local Similarity 94.7%; Pred. No. 4e+03;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 CCTTGGCGCTGGCGCCTC 20
|||||
Db 771 CCTTGGCGCTGGCGCCTC 789

RESULT 13
CNS08C9Y 83052 bp DNA linear HTG 27-JUL-2002
LOCUS Oryza sativa chromosome 12 clone OJ1410_G05, *** SEQUENCING IN
DEFINITION PROGRESS ***, 5 ordered pieces.
ACCESSION AL772420
VERSION AL772420.2 GI:22003296
KEYWORDS HTG; HTGS_PHASE2; HTGS_ACTIVEFIN.
SOURCE Oryza sativa.
ORGANISM Oryza sativa
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzaceae; Oryza.
REFERENCE 1 (bases 1 to 83052)
AUTHORS Choins, N., Orjeda, G., Cattolico, L., Demange, N., Wincker, P.,
Seguins, B., Pelletier, F., Scarpelli, C., Salanoubat, M.,
Weissenbach, J., and Quetier, F.
Oryza sativa chromosome 12 sequencing
Unpublished
2 (bases 1 to 83052)
Genoscope.
Direct Submission
Submitted (26-JUL-2002) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqrefgenoscope.cns.fr
- Web : www.genoscope.cns.fr)
On Jul 29, 2002 this sequence version replaced gi.21535766.
IMPORTANT: This sequence is unfinished and does not necessarily
represent the correct sequence.
Work on the sequence is in progress and the release of this data is
based on the understanding that the sequence may change as work
continue. The sequence may be contaminated with foreign sequence
from E.coli, yeast, vector, phage, etc.
The nucleotide sequence of this BAC clone was generated by
combining Monsanto and Genoscope sequencing data.
Contigs composition :
14140 bp contig from 1 to 14140
14899 bp contig from 14241 to 29139
7207 bp contig from 29240 to 36446
10841 bp contig from 36547 to 47387
35565 bp contig from 47488 to 83052.
* NOTE: This is a 'working draft' sequence. It currently
* consists of 5 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* provided by the submittor.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
* 1 14140: contig of 14140 bp in length
* 14141 14240: gap of 100 bp
* 14241 29139: contig of 14899 bp in length
* 29140 29239: gap of 100 bp
* 29240 36446: contig of 7207 bp in length
* 36447 36546: gap of 100 bp
* 36547 47387: contig of 10841 bp in length
* 47388 47487: gap of 100 bp
* 47488 83052: contig of 35565 bp in length.
* Location/Qualifiers
FEATURES
source 1..83052
/organism="Oryza sativa"
/cultivar="Nipponbare"
/sub_species="japonica"
/db_xref="taxon:4530"
/chromosome="12"

```

```
/db_xref="taxon:9606"
/chromosome="1"
/map="p36.21-36.33"
/clone="RPS-889P23"
/clone_lib="RPCI-5"
misc_feature 1..3331
  /note="assembly_fragment:00799"
misc_feature 3432..90998
  /note="assembly_fragment:02114"
misc_feature 91099..130106
  /note="assembly_fragment:03033.0"
misc_feature 130207..132592
  /note="assembly_fragment:03380"
BASE COUNT 26843 a 36833 c 36775 g 31841 t 300 others
ORIGIN

Query Match
Best Local Similarity 94.78; Pred. No. 1.7e+03;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 CTTTGGCGCTGCGGCTTC 20
Db 110808 CTTTGGCGCTGCGGCTTC 110826

RESULT 15
HS120G22 166518 bp DNA linear PRI 21-JUL-2000
LOCUS Human DNA sequence from clone RPI-120G22 on chromosome
DEFINITION AL031847
VERSION 1
KEYWORDS HTG.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 166518)
Wray, P.
Direct Submission
Submitted (20-JUL-2000) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk
requests: clonerequest@sanger.ac.uk
On Jul 22, 2000 this sequence version replaced gi:9368784.
During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.
This sequence has been finished according to sequence map criteria
as follows. An attempt is made to resolve all sequencing problems,
such as compressions and repeats, but not necessarily within known
annotated human repeat sequence elements (e.g. Alu). Where the
sequence is ambiguous, there is an annotation using the 'unsure'
feature key.
The following abbreviations are used to associate primary accession
numbers given in the feature table with their source databases:
Em: EMBL; Sw: SWISSPROT; Tr: TREMBL; Wp: WORMPEP; Information
on the WORMPEP database can be found at
http://www.sanger.ac.uk/projects/c_elegans/wormpep/
This sequence was generated from part of bacterial clone contigs of human
chromosome 1, constructed by the Sanger Centre Chromosome 1 Mapping
Group. Further information can be found at
http://www.sanger.ac.uk/HGP/Chr1
RPI-120G22 is from the library RPCI-1 constructed at the Roswell
Park Cancer Institute by the group of Pieter de Jong. For further
details see http://bacpac.med.buffalo.edu/
VECTOR: pcYPAC2
IMPORTANT: This sequence is not the entire insert of clone
RPI-120G22. It may be shorter because we sequence overlapping
sections only once, except for a 100 base overlap.
The true right end of clone RPI-120G22 is at 166518 in this
sequence. The true left end of clone RPS-889P23 is at 90508 in this
sequence. The true right end of clone RPI-20208 is at 100 in this
sequence.
```

FEATURES

Source	Location/Qualifiers
misc_feature	1..166518 /organism="Homo sapiens" /db_xref="taxon:9606" /chromosome="1" /map="p36.21-36.33" /clone="RPI-120G22" /clone_lib="RPCI-1" 8814..9131 /note="match: GSS: Em:AQ716455"
misc_feature	8816..9229 /note="match: GSS: Em:AQ239786"
misc_feature	10012..10541 /note="match: GSS: Em:AQ678815"
misc_feature	18389..18823 /note="match: GSS: Em:AQ270793"
misc_feature	25588..26045 /note="match: GSS: Em:AQ214655"
misc_feature	25694..25967 /note="match: GSS: Em:AQ097494"
misc_feature	complement(41087..41505) /note="match: GSS: Em:AQ213363"
misc_feature	complement(41123..41505) /note="match: GSS: Em:AQ133445"
misc_feature	45297..45591 /note="match: GSS: Em:AF046498"
misc_feature	49180..49534 /note="Single clone region. short insert library only"
misc_feature	complement(57981..58764) /note="match: GSS: Em:AF158011"
misc_feature	complement(60582..61043) /note="match: GSS: Em:AQ695221"
misc_feature	64797..64841 /note="Single clone region. short insert library only"
misc_feature	68250..68620 /note="match: STS: Em:G16281"
misc_feature	70460..70622 /note="match: STS: Em:G21080"
misc_feature	complement(74554..74895) /note="match: GSS: Em:AQ097224"
misc_feature	complement(79470..79747) /note="match: STS: Em:AF191963"
misc_feature	complement(79556..79740) /note="match: GSS: Em:AQ541496"
misc_feature	complement(79935..80319) /note="match: STS: Em:HSQA052WG1"
misc_feature	complement(85942..86372) /note="match: STS: Em:G22706"
misc_feature	complement(86963..87322) /note="match: GSS: Em:AQ191756"
misc_feature	87339..87808 /note="match: GSS: Em:AQ698012"
misc_feature	88780..89285 /note="match: GSS: Em:AQ121434"
misc_feature	94030..94434 /note="match: GSS: Em:B89150"
misc_feature	94030..94405 /note="match: GSS: Em:AQ008338"
misc_feature	94054..94464 /note="match: GSS: Em:AQ010380"
misc_feature	join(112106..112211,116809..116933,123047..123132) /note="match: GSS: Em:AQ634365"
misc_feature	complement(join(112121..112215,116809..116936)) /note="match: GSS: Em:AQ202845"
misc_feature	112126..112224 /note="match: GSS: Em:AF046385"
misc_feature	join(112145..112211,116809..116936) /note="match: GSS: Em:AQ634370"
misc_feature	complement(join(112150..112211,116809..116933,123047..123172))

us-09-270-437d-11.rge

Mon Apr 21 10:29:54 2003

```

misc_feature /note="match: GSS: Em:AQ489599"
complement(113100. 113712)
misc_feature /note="match: GSS: Em:AQ477048"
complement(113175. 113715)
misc_feature /note="match: GSS: Em:AQ476615"
113710. 114090
misc_feature /note="match: GSS: Em:AQ434236"
join(116809. 116933. 123047. 123119)
misc_feature /note="match: GSS: Em:AQ732842"
116809. 116941
misc_feature /note="match: GSS: Em:AQ357462"
116809. 116936
misc_feature /note="match: GSS: Em:AQ047297 Em:AQ351276"
join(116840. 116933. 123047. 123191)
misc_feature /note="match: STS: Em:G13224 Em:G13474"
121507. 121762
misc_feature /note="match: GSS: Em:AQ482227"
complement(123035. 123248)
misc_feature /note="match: GSS: Em:AQ779080"
123142. 123248
misc_feature /note="match: GSS: Em:AQ634365"
123192. 123356
misc_feature /note="match: STS: Em:G11177"
124453. 124584
misc_feature /note="match: GSS: Em:AQ002260"
complement(124491. 124832)
misc_feature /note="match: STS: Em:G06939"
129943. 130018
misc_feature /note="single clone region"
144105. 144499
misc_feature /note="match: GSS: Em:AQ231959"
144105. 144483
misc_feature /note="match: GSS: Em:AQ231956"
complement(148672. 149104)
misc_feature /note="match: GSS: Em:AQ672446"
152031. 152510
misc_feature /note="match: GSS: Em:AQ234353"
152031. 152510
BASE COUNT 34311 a 45982 c 47124 g 39101 t
ORIGIN

```

Query Match 79.1%; Score 17.4; DB 9; Length 166518;
 Best Local Similarity 94.7%; Pred. No. 1.6e+03;
 Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2 CCTTGGCGCTGCGGCTC 20
 |||||
 Db 110222 CCTTGGCGCTGCGGCTC 110240

Search completed: April 18, 2003, 06:20:58
 Job time : 508 secs

GenCore version 5.1.4 p5 4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 18, 2003, 04:12:35 ; Search time 98 Seconds
(without alignments)
505.551 Million cell updates/sec

Title: US-09-270-437D-11

Perfect score: 22
Sequence: 1 tctctggcgctgcggcctcag 22

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 112599159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

N_Geneseq_101002.*
1: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*
2: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*
3: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*
4: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*
5: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*
6: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*
7: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*
8: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*
9: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*
10: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*
11: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*
12: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*
13: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*
14: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*
15: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*
16: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*
17: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*
18: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*
19: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*
20: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*
21: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*
22: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*
23: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*
24: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	22	100.0	22	21	AAZ36157
2	22	100.0	1945	21	AAZ36153
3	17.8	80.9	139	21	AAZ37338
4	17.8	80.9	1711	22	AAH07034
5	17.8	80.9	1785	23	AAH08110
6	17.8	80.9	1785	23	AAH08110
7	17.8	80.9	2082	22	AAH14706
8	17.2	78.2	464	22	AAH63986
9	17.2	78.2	1217	23	ABL29179

C 10	17.2	78.2	3408	23	ABL29178	Drosophila melanog
C 11	16.8	76.4	462	22	AAZ36157	Human breast cancer
C 12	16.8	76.4	462	22	AAZ36157	Human breast cancer
C 13	16.8	76.4	462	24	AAZ36157	Human breast cancer
C 14	16.8	76.4	10461	20	AAZ36157	Polynucleotide seq
C 15	16.8	76.4	13308	24	AAZ36157	Human enolase 3 (be
C 16	16.4	74.5	151	20	AAZ36157	Human enolase 3 (be
C 17	16.4	74.5	333	22	AAZ36157	Human enolase 3 (be
C 18	16.4	74.5	333	22	AAZ36157	Human enolase 3 (be
C 19	16.4	74.5	762	22	AAZ36157	Human enolase 3 (be
C 20	16.4	74.5	762	22	AAZ36157	Human enolase 3 (be
C 21	16.4	74.5	1580	22	AAZ36157	Human enolase 3 (be
C 22	16.4	74.5	1620	21	AAZ36157	Human enolase 3 (be
C 23	16.4	74.5	1759	21	AAZ36157	Human enolase 3 (be
C 24	16.4	74.5	1759	22	AAZ36157	Human enolase 3 (be
C 25	16.4	74.5	1759	22	AAZ36157	Human enolase 3 (be
C 26	16.4	74.5	1875	22	AAZ36157	Human enolase 3 (be
C 27	16.4	74.5	1885	22	AAZ36157	Human enolase 3 (be
C 28	16.4	74.5	1885	22	AAZ36157	Human enolase 3 (be
C 29	16.4	74.5	1931	20	AAZ36157	Human enolase 3 (be
C 30	16.4	74.5	1932	20	AAZ36157	Human enolase 3 (be
C 31	16.4	74.5	1940	22	AAZ36157	Human enolase 3 (be
C 32	16.4	74.5	1940	22	AAZ36157	Human enolase 3 (be
C 33	16.4	74.5	1974	22	AAZ36157	Human enolase 3 (be
C 34	16.4	74.5	12585	22	AAZ36157	Human enolase 3 (be
C 35	16.2	73.6	208	21	AAZ36157	Human enolase 3 (be
C 36	16.2	73.6	442	21	AAZ36157	Human enolase 3 (be
C 37	16.2	73.6	442	21	AAZ36157	Human enolase 3 (be
C 38	16.2	73.6	605	22	AAZ36157	Human enolase 3 (be
C 39	16.2	73.6	622	22	AAZ36157	Human enolase 3 (be
C 40	16.2	73.6	789	22	AAZ36157	Human enolase 3 (be
C 41	16.2	73.6	1368	23	AAZ36157	Human enolase 3 (be
C 42	16.2	73.6	2072	22	AAZ36157	Human enolase 3 (be
C 43	16.2	73.6	2409	19	AAZ36157	Human enolase 3 (be
C 44	16.2	73.6	2409	19	AAZ36157	Human enolase 3 (be
C 45	16.2	73.6	2583	24	AAZ36157	Human enolase 3 (be

ALIGNMENTS

RESULT 1
AAZ36157
ID AAZ36157 standard; DNA; 22 BP.
XX
AC AAZ36157;
XX

DT 11-FEB-2000 (first entry)
XX
DE PCR primer for DNA encoding cancer associated antigen KOC-2.
XX
KW Cancer associated antigen; KOC-1; cancer; vaccine; CT7; PCR primer; ss.

OS Synthetic.
OS Homo sapiens.

PN WO9954738-A1.
XX

PD 28-OCT-1999.
XX

PF 16-MAR-1999; 99WO-US05766.
XX

PR 17-APR-1998; 98US-0061709.
XX

XX (LUDW-) LUDWIG INST CANCER RES.
XX

PI Chen Y, Gure A, Tsang S, Stockert E, Jager E, Knuth A, Old LJ;
XX

DR WPI; 2000-013284/01.
XX

PT Nucleotides representing cancer-associated genes, used to develop
XX products for the diagnosis, monitoring and treatment of cancers

Mon Apr 21 10:29:56 2003

or vaccine composition with an adjuvant, e.g. a cytokine, a saponin, or granulocyte macrophage-colony stimulating factor (GM-CSF).

CC Sequence 1946 BP; 502 A; 528 C; 553 G; 358 T; 5 other;
 XX
 SQ
 Query Match 100.0%; Score 22; DB 21; Length 1946;
 Best Local Similarity 100.0%; Pred. No. 2.5;
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCCTTGGCGCTGGCGCTCAG 22
 |||||
 DB 38 TCCTTGGCGCTGGCGCTCAG 59
 |||||

RESULT 3
 AAC27338/c
 ID AAC27338 standard; cDNA; 139 BP.
 XX
 AC AAC27338;
 XX
 DT 06-OCT-2000 (first entry)
 DE Human secreted protein 5' EST, SEQ ID NO: 31413.
 XX
 XX Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
 KW gene therapy; chromosome mapping; ss.
 KW Homo sapiens.
 OS
 XX
 FN EP1033401-A2.
 XX
 XX 06-SEP-2000.
 PD
 XX 21-FEB-2000; 2000EP-0200610.
 PF
 XX 26-FEB-1999; 99US-0122487.
 PR
 XX (GIST) GENSET.
 XX
 XX Dumas Milne Edwards J, Duclert A, Giordano J;
 XX WPI; 2000-500381/45.
 DR
 XX New nucleic acid that is a 5' expressed sequence tag (5' EST) for
 PT obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for
 PT diagnostic, forensic, gene therapy and chromosome mapping procedures -
 XX Claim 1; SEQ ID 31413; 71pp + CD-ROM; English.

The present sequence is one of a large number of 5' ESTs derived from mRNAs encoding secreted proteins. No ORF has yet been conclusively identified within the present sequence. The 5' ESTs were prepared from total human RNAs or polyA+ RNAs derived from 30 different tissues. EST sequences usually correspond mainly to the 3' untranslated region (UTR) of the mRNA because they are often obtained from oligo-dT primed cDNA libraries. Such ESTs are not well suited for isolating cDNA sequences derived from the 5' ends of mRNAs and even in those cases where longer cDNA sequences have been obtained, the full 5' UTR is rarely included. 5' ESTs are derived from mRNAs with intact 5' ends and can therefore be used to obtain full length cDNAs and genomic DNAs. 5' ESTs are also used in diagnostic, forensic, gene therapy and chromosome mapping procedures. They are used to obtain upstream regulatory sequences and to design expression and secretion vectors.

XX Sequence 139 BP; 26 A; 44 C; 54 G; 14 T; 1 other;
 SQ
 Query Match 80.9%; Score 17.8; DB 21; Length 139;
 Best Local Similarity 90.5%; Pred. No. 1.5e+02;
 Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 CCTTGGCGCTGGCGCTCAG 22
 |||||
 DB 50 CCTTGGCGCTGGCGCTCAG 30
 |||||

PS Claim 108; Page 13; 44pp; English.
 XX
 CC PCR primers AAZ36157-58 were used to amplify a cancer associated antigen
 CC gene designated KOC-2. The specification also describes a cancer
 CC associated antigen designated CT7. The CT7 polynucleotide was isolated
 CC from SK-MEL-37 melanoma cells. The polypeptide has some homology with
 CC MAGE-10, limited to about 210 carboxy terminal amino acids. The amino
 CC terminal of the protein has a repetitive pattern, with repeats rich in
 CC serine, proline, glutamine and leucine, and an almost invariable core of
 CC the peptide given in AAY43877. The CT7 polypeptide can be processed to
 CC peptides which provoke lysis by cytolytic T cells. The polynucleotides
 CC and polypeptides can be used for treating a cancerous condition and
 CC screening for or diagnosing cancerous conditions. The cancer associated
 CC antigens can be used as an immunogenic or vaccine composition with an
 CC adjuvant, e.g. a cytokine, a saponin, or granulocyte macrophage-colony
 CC stimulating factor (GM-CSF).
 XX
 SQ Sequence 22 BP; 1 A; 9 C; 7 G; 5 T; 0 other;

Query Match 100.0%; Score 22; DB 21; Length 22;
 Best Local Similarity 100.0%; Pred. No. 2.5;
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCCTTGGCGCTGGCGCTCAG 22
 |||||
 DB 1 TCCTTGGCGCTGGCGCTCAG 22
 |||||

RESULT 2
 AAZ36153
 ID AAZ36153 standard; DNA; 1946 BP.
 XX
 AC AAZ36153;
 XX
 DT 11-FEB-2000 (first entry)
 DE An alternative form of DNA encoding cancer associated antigen KOC-2.
 XX
 XX Cancer associated antigen; KOC-2; cancer; vaccine; CT7; ss.
 KW Homo sapiens.
 OS
 XX
 FN WO9954738-A1.
 XX
 XX 28-OCT-1999.
 PD
 XX 16-MAR-1999; 99WO-US05766.
 PF
 XX 17-APR-1998; 98US-0061709.
 PR
 XX (LUDW-) LUDWIG INST CANCER RES.
 XX
 XX Chen Y, Gure A, Tsang S, Stockert E, Jager E, Knuth A, Old LJ;
 XX WPI; 2000-013284/01.
 DR
 XX Nucleotides representing cancer-associated genes, used to develop
 PT products for the diagnosis, monitoring and treatment of cancers -
 XX Claim 55; Page 42; 44pp; English.

The present sequence represents an alternative form of a cancer associated antigen gene designated KOC-2. The specification also describes a cancer associated antigen designated CT7. The CT7 polynucleotide has some homology with MAGE-10, limited to about 210 carboxy terminal amino acids. The amino terminal of the protein has a repetitive pattern, with repeats rich in serine, proline, glutamine and leucine, and an almost invariable core of the peptide given in AAY43877. The CT7 polypeptide can be processed to peptides which provoke lysis by cytolytic T cells. The polynucleotides and polypeptides can be used for treating a cancerous condition and screening for or diagnosing cancerous conditions. The cancer associated antigens can be used as an immunogenic

RESULT 4
AAH07034/C
ID AAH07034 standard; cDNA; 711 BP.
XX
AC AAH07034;
XX
DT 26-JUN-2001 (first entry)
XX
DE Human cDNA clone (5'-primer) SEQ ID NO:3869.
XX
KW Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
XX
OS Homo sapiens.
XX
PN EP1074617-A2.
XX
PD 07-FEB-2001.
XX
PF 28-JUL-2000; 2000EP-0116126.
XX
PR 29-JUL-1999; 99JP-0248036.
PR 27-AUG-1999; 99JP-0300253.
PR 11-JAN-2000; 2000JP-0118776.
PR 02-MAY-2000; 2000JP-0183767.
PR 09-JUN-2000; 2000JP-0241899.
XX
PA (HELI-) HELIX RES INST.
XX
PI Ota T, Isogai T, Nishikawa T, Hayaashi K, Saito K, Yamamoto J;
PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX
DR WPI; 2001-318749/34.
XX
PT Primer sets for synthesizing polynucleotides, particularly the 5602
PT full-length cDNAs defined in the specification, and for the detection
PT and/or diagnosis of the abnormality of the proteins encoded by the
PT full-length cDNAs -
XX
PS Claim 1; SEQ ID 3869; 2537pp + CD ROM; English.
XX
CC The present invention describes primer sets for synthesising 5602
CC full-length cDNAs defined in the specification. Where a primer set
CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary
CC to the complementary strand of a polynucleotide which comprises one of
CC the 5602 nucleotide sequences defined in the specification, where the
CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC of an oligonucleotide comprising a sequence complementary to the
CC complementary strand of a polynucleotide which comprises a 5'-end
CC sequence and an oligonucleotide comprising a sequence complementary to a
CC polynucleotide which comprises a 3'-end sequence, where the
CC oligonucleotide comprises at least 15 nucleotides and the combination of
CC the 5'-end sequence/3'-end sequence is selected from those defined in
CC the specification. The primer sets can be used in antisense therapy and
CC in gene therapy. The primers are useful for synthesising polynucleotides,
CC detection and/or diagnosis of the abnormality of the proteins encoded by
CC the full-length cDNAs. The primers are also useful for the
CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
CC AAH13633 to AAH18742 represent human cDNA sequences; AAH92446 to
CC AAH95893 represent human amino acid sequences; and AAH13629 to AAH13632
CC represent oligonucleotides, all of which are used in the exemplification
CC of the present invention.
XX
SQ Sequence 711 BP; 149 A; 239 C; 209 G; 111 T; 3 other;
Query Match 80.9%; Score 17.8; DB 22; Length 711;
Best Local Similarity 90.5%; Pred. No. 1.5e+02;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 2 CCTTGGCGCTGCGCCCTCAG 22
|||||

Db 39 CCTTGGCGCTGCGCCCTCAG 19
RESULT 5
AAS68110/C
ID AAS68110 standard; cDNA; 1785 BP.
XX
AC AAS68110;
XX
DT 13-FEB-2002 (first entry)
XX
DE DNA encoding novel human diagnostic protein #3914.
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
OS Homo sapiens.
XX
PN WO200175067-A2.
XX
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US08631.
XX
PR 31-MAR-2000; 2000US-0540217.
PR 23-AUG-2000; 2000US-0649167.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Drmanac RT, Liu C, Tang YT;
XX
DR WPI; 2001-639362/73.
XX
DR P-PSDB; ABG03923.
XX
PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity -
XX
PS Claim 1; SEQ ID NO 3914; 103pp; English.
XX
CC The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC mapping and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful in treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. AAS64197-AAS94564 represent novel human
CC diagnostic coding sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 1785 BP; 416 A; 497 C; 517 G; 355 T; 0 other;
Query Match 80.9%; Score 17.8; DB 23; Length 1785;
Best Local Similarity 90.5%; Pred. No. 1.5e+02;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 2 CCTTGGCGCTGCGCCCTCAG 22
|||||

AAH14706/c
ID AAH14706 standard; cDNA; 2082 BP.
XX
XX AC AAH14706;
XX
XX 26-JUN-2001 (first entry)
DT
XX Human cDNA sequence SEQ ID NO:12421.
DE
XX Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
XX
XX Homo sapiens.
OS
XX EF1074617-A2.
DN
XX 07-FEB-2001.
PD
XX 28-JUL-2000; 2000EP-0116126.
XX
XX 29-JUL-1999; 99JP-0248036.
XX
XX 27-AUG-1999; 98JP-0300253.
PR
XX 11-JAN-2000; 2000JP-0118776.
PR
XX 02-MAY-2000; 2000JP-0183767.
PR
XX 09-JUN-2000; 2000JP-0241899.
PR
XX
XX (HELI-) HELIX RES INST.
PA
XX Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
PI
PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX
XX WPI; 2001-318749/34.
DR
XX
XX Primer sets for synthesizing polynucleotides, particularly the 5602
PT full-length cDNAs defined in the specification, and for the detection
PT and/or diagnosis of the abnormality of the proteins encoded by the
PT full-length cDNAs -
PT
XX Claim 8; SEQ ID 12421; 2537pp + CD ROM; English.
PS
XX The present invention describes primer sets for synthesising 5602
CC full-length cDNAs defined in the specification. Where a primer set
CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary
CC to the complementary strand of a polynucleotide which comprises one of
CC the 5602 nucleotide sequences defined in the specification, where the
CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC of an oligonucleotide comprising a sequence complementary to the
CC complementary strand of a polynucleotide which comprises a 5'-end
CC sequence and an oligonucleotide comprising a sequence complementary to a
CC polynucleotide which comprises a 3'-end sequence, where the
CC polynucleotide comprises at least 15 nucleotides and the combination of
CC the 5'-end sequence/3'-end sequence is selected from those defined in
CC the specification. The primer sets can be used in antisense therapy and
CC in gene therapy. The primers are useful for synthesising polynucleotides,
CC particularly full-length cDNAs. The primers are also useful for the
CC detection and/or diagnosis of the abnormality of the proteins encoded by
CC the full-length cDNAs. The primers allow obtaining of the full-length
CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to
CC AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632
CC represent oligonucleotides, all of which are used in the exemplification
CC of the present invention.
XX
XX Sequence 2082 BP: 416 A: 684 C: 607 G: 375 T: 0 other;

	Query Match	80.9%;	Score 17.8;	DB 22;	Length 2082;
	Best Local Similarity	90.5%;	Pred. No. 1.5e+02;		
	Matches 19;	Conservative	0;	Mismatches 2;	Indels 0;
					Gaps 0;
QY	2	CCTTGC	CGCGCTG	CGGCCTCAG	22
Db	39	CCTTGC	CGCGCTG	CAGCCCCAG	19

RESULT 6
 AAS73402
 ID AAS73402 standard; cDNA; 1785 BP.
 XX
 XX AAS73402;
 XX
 XX
 DT 13-FEB-2002 (first entry)
 XX
 XX DNA encoding novel human diagnostic protein #9206.
 DE
 XX
 XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
 KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
 KW
 XX
 XX Homo sapiens.
 OS
 XX
 XX WO200175067-A2.
 PN
 XX
 PD 11-OCT-2001.
 XX
 PD 30-MAR-2001; 2001WO-US08631.
 PF
 PF 31-MAR-2000; 2000US-0540217.
 PR 23-AUG-2000; 2000US-0649167.
 PR
 XX (HYSE-) HYSEQ INC.
 PA
 PA Drmanac RT, Liu C, Tang YT;
 PI
 PI WPI; 2001-639362/73.
 DR P-PSDB; ABG09215.
 DR
 DR
 XX New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity -
 XX
 XX Claim 1; SEQ ID No 9206; 103pp; English.
 PS
 PS
 CC The invention relates to isolated polynucleotide (I) and
 CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
 CC and gene mapping, and in recombinant production of (II). The
 CC polynucleotides are also used in diagnostics as expressed sequence tags
 CC for identifying expressed genes. (I) is useful in gene therapy techniques
 CC to restore normal activity of (II) or to treat disease states involving
 CC (II). (II) is useful for generating antibodies against it, detecting or
 CC quantitating a polypeptide in tissue, as molecular weight markers and as
 CC a food supplement. (II) and its binding partners are useful in medical
 CC imaging of sites expressing (II). (I) and (II) are useful for treating
 CC disorders involving aberrant protein expression or biological activity.
 CC The polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. AAS64197-AAS94564 represent novel human
 CC diagnostic coding sequences of the invention.
 CC Note: The sequence data for this patent did not appear in the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 CC
 XX Sequence 1785 BP: 355 A; 517 C; 497 G; 416 T; 0 other;

Query Match	80.8%	Score	17.8	DB	23	Length	1785
Best Local Similarity	90.5%	Pred. No.	1.5e+02				
Matches	19	Conservative	0	Mismatches	2	Indels	0
Gaps	0						

RESULT 7

RESULT 8
AAK63986/c
ID AAK63986 standard; cDNA; 464 BP.
XX
AC AAK63986;
XX
DT 06-NOV-2001 (first entry)
XX
DE Human immune/haematopoietic antigen encoding cDNA SEQ ID NO:9046.
XX
KW Human; Immune; haematopoietic; immune/haematopoietic antigen; cancer;
cytostatic; gene therapy; vaccine; metastasis; ss.
XX Homo sapiens.
OS
XX
PN WO200157182-A2.
XX
PD
XX
XX 09-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US01354.
XX
XX 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216680.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 18-AUG-2000; 2000US-0225759.
PR 22-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226868.
PR 23-AUG-2000; 2000US-0227182.
PR 30-AUG-2000; 2000US-0227009.
PR 01-SEP-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-02331968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239335.
PR 13-OCT-2000; 2000US-0239337.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 01-DEC-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.

PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
PA (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Barash SC, Ruben SM;
PI WPI; 2001-483426/52.
XX P-PSDB; AAM91205.
XX
XX Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
PT useful for preventing, diagnosing and/or treating cancers and
PT metastasis -
XX
XX Claim 1; SEQ ID NO 9046; 3071pp + Sequence Listing; English.
PS
XX AAK54951 to AAK64702 encode the human immune/hematopoietic antigen (I)
XX amino acid sequences given in AAM82170 to AAM91921. (I) have cytostatic
XX activity, and can be used in gene therapy and vaccine production. (I)
XX proteins and polynucleotides may be used in the prevention, diagnosis and
XX treatment of diseases associated with inappropriate (I) expression. For
XX example, they may be used to treat disorders associated with decreased
XX expression by rectifying mutations or deletions in a patient's genome
XX that affect the activity of (I) by expressing inactive proteins or to
XX supplement the patient's own production of (I). Additionally, (I)
XX polynucleotides may be used to produce the secreted (I), by inserting
XX the nucleic acids into a host cell and culturing the cell to express the
XX protein. (I) proteins and polynucleotides may be used to prevent,
XX diagnose and treat immune/hematopoietic-related diseases, especially
XX cancers and cancer metastases of hematopoietic-derived cells. AAK4703
XX to AAK87694 represent human immune/hematopoietic antigen genomic
XX sequences from the present invention. AAK54942 to AAK54950 and AAM82169
XX represent sequences used in the exemplification of the present invention.
XX
XX Sequence 464 BP; 131 A; 115 C; 110 G; 107 T; 1 other;
SQ
Query Match 78.2%; Score 17.2; DB 22; Length 464;
Best Local Similarity 86.4%; Pred. No. 2.7e+02;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 TCCTTGGCGCTGGCGCTCAG 22
Db 200 TCCTTGGCGCTGGCGCTCTG 179
RESULT 9
ABL29179/c
ID ABL29179 standard; DNA; 1217 BP.
XX
XX ABL29179;
AC
XX 26-MAR-2002 (first entry)
DT
XX Drosophila melanogaster genomic polynucleotide SEQ ID NO 39010.
DE
XX Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical; gene; ds.
XX
XX Drosophila melanogaster.
OS
XX WO200171042-A2.
PN
XX 27-SEP-2001.
PD
XX 23-MAR-2001; 2001WO-US09231.
PF
XX 23-MAR-2000; 2000US-191637P.
PR 11-JUL-2000; 2000US-0614150.
PR
XX (PEXE) PE CORP NY.
PA
XX Venter JC, Adams M, Li PWD, Myers EW;
PI WPI; 2001-656860/75.
XX
XX New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions -
PT

PF 23-MAR-2001; 2001WO-US09231.
XX
PR 23-MAR-2000; 2000US-191637P.
PR 11-JUL-2000; 2000US-0614150.
XX
XX (PEXE) PE CORP NY.
XX
XX Venter JC, Adams M, Li PWD, Myers EW;
PI WPI; 2001-656860/75.
XX
XX New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions -
XX
XX Claim 1; SEQ ID NO 39010; 21pp + Sequence Listing; English.
PS
XX The invention relates to an isolated nucleic acid detection reagent
XX capable of detecting 1000 or more genes from Drosophila. The invention is
XX useful in developmental biology and in elucidating cell signalling and
XX cell-cell interactions in higher eukaryotes for the development of
XX insecticides, therapeutics and pharmaceutical drugs. The invention
XX discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
XX sequences (ABL01840-ABL16175) and the encoded proteins
XX (ABBS7737-ABBY2072).
XX The sequence data for this patent did not form part of the printed
XX specification, but was obtained in electronic format directly from WIPO
XX at ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 1217 BP; 301 A; 328 C; 358 G; 230 T; 0 other;
SQ
Query Match 78.2%; Score 17.2; DB 23; Length 1217;
Best Local Similarity 86.4%; Pred. No. 2.7e+02;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 TCCTTGGCGCTGGCGCTCAG 22
Db 1066 TCCTTGGCGCTGGCGCTCAG 1045
RESULT 10
ABL29178/c
ID ABL29178 standard; DNA; 3408 BP.
XX
XX ABL29178;
AC
XX 26-MAR-2002 (first entry)
DT
XX Drosophila melanogaster genomic polynucleotide SEQ ID NO 39007.
DE
XX Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical; gene; ds.
XX
XX Drosophila melanogaster.
OS
XX WO200171042-A2.
PN
XX 27-SEP-2001.
PD
XX 23-MAR-2001; 2001WO-US09231.
PF
XX 23-MAR-2000; 2000US-191637P.
PR 11-JUL-2000; 2000US-0614150.
PR
XX (PEXE) PE CORP NY.
PA
XX Venter JC, Adams M, Li PWD, Myers EW;
PI WPI; 2001-656860/75.
XX
XX New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions -
PT

XX PS Claim 1; SEQ ID NO 39007; 21pp + Sequence Listing; English.

XX CC The invention relates to an isolated nucleic acid detection reagent

CC capable of detecting 1000 or more genes from *Drosophila*. The invention is

CC useful in developmental biology and in elucidating cell signalling and

CC cell-cell interactions in higher eukaryotes for the development of

CC insecticides, therapeutics and pharmaceutical drugs. The invention of

CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA

CC sequences (ABL01840-ABL116175) and the encoded proteins

CC (ABB57737-ABB72072).

CC The sequence data for this patent did not form part of the printed

CC specification, but was obtained in electronic format directly from WIPO

CC at ftp.wipo.int/pub/published_pct_sequences.

XX SQ Sequence 3408 BP; 945 A; 779 C; 790 G; 894 T; 0 other;

Query Match 78.2%; Score 17.2; DB 23; Length 3408;

Best Local Similarity 86.4%; Pred. No. 2.6e+02;

Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 TCCTTGGCGCTGCGGCTCAG 22

Db 2257 TCCTTGGCGCTGCGGCTCAG 2236

RESULT 11

AA547110/C

ID AAS47110 standard; cDNA; 462 BP.

XX AC AAS47110;

XX DT 18-DEC-2001 (first entry)

XX DE Human breast cancer cDNA clone 14382.

XX KW Human; ss; breast cancer protein; tumour; cancer; cytostatic;

XX KW gene therapy.

XX OS Homo sapiens.

XX PN WO200179286-A2.

XX PD 25-OCT-2001.

XX PF 12-APR-2001; 2001WO-US12164.

XX PR 17-APR-2000; 2000US-0551621.

PR 08-JUN-2000; 2000US-0590751.

PR 22-JUN-2000; 2000US-0604287.

PR 20-JUL-2000; 2000US-0620405.

XX PA (CORI-) CORIXA CORP.

XX PI Jiang Y, Dillon DC, Mitcham JL, Xu J, Harlocker SL, Hepler WT;

XX DR WPI; 2001-611721/70.

XX PT Breast Tumour Proteins and nucleic acids useful for the prevention,

PT diagnosis and treatment of breast cancer -

XX Claim 37; Page 193; 297pp; English.

XX CC The invention relates to isolated breast tumour proteins and

CC nucleic acids that encode them, including immunogenic fragments of the

CC proteins. Also included are expression vectors expressing the

CC proteins, transformed cells and antibodies raised against the proteins or

CC an antigen presenting cell expressing the protein. The proteins and

CC nucleic acids may be used in the prevention, diagnosis and treatment of

CC diseases associated with inappropriate breast tumour protein expression,

CC i.e. breast tumours and breast cancer e.g. by gene therapy. The nucleic

CC acids and their complements may also be used as DNA probes in diagnostic

CC assays to detect and quantitate the presence of similar nucleic acids in

CC samples, and therefore which patients may be in need of restorative

CC therapy. The proteins, nucleic acids and antibodies may be used in assays

CC to identify modulators (e.g. antagonists) of breast tumour protein

CC expression and activity. The antibodies and antagonists may also be used

CC to down regulate expression and activity. The antibodies may also be used

CC as diagnostic agents for detecting the presence of the proteins in

CC samples (e.g. by enzyme linked immunosorbent assay (ELISA)) and in other

CC immuno-purification diagnostic techniques. The present sequence is

CC a cDNA from a breast tumour cDNA library isolated by subtractive

CC hybridisation against a normal breast cDNA library.

XX SQ Sequence 462 BP; 124 A; 108 C; 118 G; 99 T; 13 other;

Query Match 76.4%; Score 16.8; DB 22; Length 462;

Best Local Similarity 90.0%; Pred. No. 4e+02;

Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 TCCTTGGCGCTGCGGCTC 20

Db 91 TCTTTGCTCGCTGCGGCTC 72

RESULT 12

AAF17680/C

ID AAF17680 standard; cDNA; 462 BP.

XX AC AAF17680;

XX DT 13-MAR-2001 (first entry)

XX DE Human breast cancer associated 14382 coding sequence.

XX KW Human; breast cancer associated gene; vaccine; diagnosis; therapy; ss.

XX OS Homo sapiens.

XX PN WO200060076-A2.

XX PD 12-OCT-2000.

XX PF 15-FEB-2000; 2000WO-US05308.

XX PR 02-APR-1999; 99US-0285480.

PR 23-JUN-1999; 99US-0339338.

PR 02-SEP-1999; 99US-0389681.

PR 03-NOV-1999; 99US-0433826.

XX PA (CORI-) CORIXA CORP.

XX PI Yuqiu J, Dillon DC, Mitcham JL, Xu J, Harlocker SL;

XX DR WPI; 2001-122627/13.

XX PT An isolated polypeptide useful for the treatment and diagnosis of

PT tumours e.g. breast cancer comprises at least an immunogenic portion of

PT a breast tumor protein -

XX Claim 66; Page 138; 238pp; English.

XX CC The present invention provides the coding sequences and some protein

CC sequences of proteins associated with breast cancer in humans. These

CC sequences can be used in the diagnosis and treatment of cancers,

CC particularly breast tumours.

XX SQ Sequence 462 BP; 124 A; 108 C; 118 G; 99 T; 13 other;

Query Match 76.4%; Score 16.8; DB 22; Length 462;

Best Local Similarity 90.0%; Pred. No. 4e+02;

Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 TCCTTGGCGCTGCGGCTC 20

Db 91 TCTTTGCTCGCTGCGGCTC 72

```
RESULT 13
ABK95145/c
ID ABK95145 standard; cDNA; 462 BP.
XX
XX AC ABK95145;
XX
XX DT 24-SEP-2002 (first entry)
XX
XX DE Human breast tumour cDNA 14382.
XX
XX KW Human; ss; breast cancer; cytostatic; gene therapy; SYN22A12; SYN22A2;
XX B723P; B726P.
XX
XX OS Homo sapiens.
XX
XX PN US6387697-B1.
XX
XX PD 14-MAY-2002.
XX
XX PF 28-DEC-1998; 98US-0222575.
XX
XX PR 28-DEC-1998; 98US-0222575.
XX
XX PA (CORI-) CORIXA CORP.
XX
XX PI Yuqiu J, Dillon DC, Mitcham JL, Xu J;
XX WPI; 2002-478446/51.
XX
XX PT New nucleic acids encoding breast cancer antigens SYN22A12, SYN22A2,
XX B723P and B726P, useful for the prevention, diagnosis and treatment of
XX breast cancer.
XX
XX PS Example 2; Column 147-148; 82pp; English.
XX
XX CC The invention relates to isolated polynucleotides: (I)-(IV) encoding
XX breast cancer antigens SYN22A12, SYN22A2, B723P and B726P (4 of 172 cDNA
XX sequences included in the specification). Also included are isolated
XX cDNAs comprising sequences with 90 % identity to (I)-(IV), an expression
XX vector comprising the cDNAs, a host cell transformed with the expression
XX vector and a diagnostic kit comprising 2 oligonucleotide primers or
XX probes, where 1 of the oligonucleotide primers or probes is specific
XX for (I)-(IV) and isolated cDNAs comprising sequences with 90 %
XX identity to (I)-(IV), and that is 10 nucleotides in length.
XX These polynucleotide sequences (I) to (IV) and the SYN22A12, SYN22A2,
XX B723P and B726P peptides they encode can be used as diagnostic
XX markers, as models for the development of human therapeutic targets,
XX aid in the identification of therapeutic proteins, and serve as targets
XX for the development of human therapeutic agents that may be used for
XX the treatment of breast cancer. In particular, they may be transfected
XX (using a vector) into a host cell that may be cultured to express the
XX breast cancer antigens. These antigens may then be used in the production
XX of antibodies against SYN22A12, SYN22A2, B723P and B726P, which in turn
XX may be used as immunoassay reagent and therapeutics for the diagnosis and
XX treatment of breast cancer. Short 10 nucleotide sequences of (I) to (IV)
XX may be used as primers and probes in hybridisation and amplification
XX reactions for the detection of SYN22A12, SYN22A2, B723P and B726P in the
XX diagnosis of breast cancer. The present sequence is a breast cancer
XX specific cDNA of the invention.
XX
XX SQ Sequence 462 BP; 124 A; 108 C; 118 G; 99 T; 13 other;

Query Match 76.4%; Score 16.8; DB 24; Length 462;
Best Local Similarity 90.0%; Pred. No. 4e+02; Indels 0; Gaps 0;
Matches 18; Conservative 0; Mismatches 2;

Qy 1 TCCTTGGCGCTGGCGCTC 20
Db 91 TCTTTGCTGGCTGGCGCTC 72

RESULT 14
AA20553/c
ID AA20553 standard; DNA; 10461 BP.
XX
XX AC AA20553;
XX
XX DT 05-MAY-1999 (first entry)
XX
XX DE Polynucleotide sequence from the genome of Treponema pallidum.
XX
XX KW Treponema pallidum infection; syphilis; Borrelia infection; animal;
XX enzyme production; ds.
XX
XX OS Treponema pallidum.
XX
XX PN WO9859034-A2.
XX
XX PD 30-DEC-1998.
XX
XX PF 23-JUN-1998; 98WO-US13041.
XX
XX PR 24-JUN-1997; 97US-0050667.
XX
XX PA (HUMA-) HUMAN GENOME SCI INC.
XX
XX PI Fraser CM;
XX WPI; 1999-081273/07.
XX
XX PT New isolated Treponema pallidum nucleic acids - used to develop
XX products for the detection, diagnosis, characterisation, prevention
XX and therapy of T. pallidum infections, particularly syphilis
XX
XX PS Claim 1; Page 491-497; 1150pp; English.
XX
XX CC AA20500-21243 represent polynucleotide sequences from the genome of
XX Treponema pallidum. The sequences can be used for detection,
XX diagnosis, characterisation, prevention and therapy for T. pallidum
XX infections, particularly syphilis. They can also be used for detecting
XX diseases related to Borrelia infections in animals, and for the
XX production of biosynthetic products such as enzymes.
XX
XX SQ Sequence 10461 BP; 2325 A; 3411 C; 2493 G; 2213 T; 19 other;

Query Match 76.4%; Score 16.8; DB 20; Length 10461;
Best Local Similarity 90.0%; Pred. No. 3.9e+02; Indels 0; Gaps 0;
Matches 18; Conservative 0; Mismatches 2;

Qy 3 CTTGGCGCTGGCGCTCAG 22
Db 6012 CTTGGCGCTGGCGCTCAG 5993

RESULT 15
AAK98531
ID AAK98531 standard; DNA; 13308 BP.
XX
XX AC AAK98531;
XX
XX DT 16-APR-2002 (first entry)
XX
XX DE Human enolase 3 (beta, muscle) isogene.
XX
XX KW Human; enolase 3 (beta, muscle); ENO3; single nucleotide polymorphism;
XX SNP; haplotype analysis; isogene; gene; ds.
XX
XX OS Homo sapiens.
XX
XX FH Key Location/Qualifiers
XX CDS 4092..9309
XX /*tag= a
XX /*product= "ENO3"
XX /*note= "this sequence contains introns"
```



```
FT allele replace(3640,C)
FT *tag= b
FT exon 4092..4176
FT *tag= c
FT intron /number= "1"
FT 4177..5056
FT *tag= d
FT /number= "1"
FT allele replace(5039,C)
FT *tag= e
FT exon 5057..5152
FT *tag= f
FT intron /number= "2"
FT 5153..5312
FT *tag= g
FT /number= "2"
FT exon 5313..5370
FT *tag= h
FT /number= "3"
FT allele replace(5343,G)
FT *tag= i
FT allele replace(5357,C)
FT *tag= j
FT allele replace(5366,C)
FT *tag= k
FT intron 5371..5532
FT *tag= l
FT exon /number= "3"
FT 5533..5603
FT *tag= m
FT /number= "4"
FT allele replace(5542,C)
FT *tag= n
FT allele replace(5547,C)
FT *tag= o
FT intron 5604..5973
FT *tag= p
FT /number= "4"
FT exon 5974..6107
FT *tag= q
FT intron /number= "5"
FT 6108..7336
FT *tag= r
FT /number= "5"
FT allele replace(7297,G)
FT *tag= s
FT exon 7337..7559
FT *tag= t
FT intron /number= "6"
FT 7560..7668
FT *tag= u
FT exon /number= "6"
FT 7669..7866
FT *tag= v
FT /number= "7"
FT allele replace(7790,G)
FT *tag= w
FT intron 7867..8203
FT *tag= x
FT /number= "7"
FT allele replace(8181,G)
FT *tag= y
FT exon 8204..8405
FT *tag= z
FT /number= "8"
FT allele replace(8332,G)
FT *tag= aa
FT intron 8406..8834
FT *tag= ab
FT /number= "8"
FT allele replace(8423,C)
FT *tag= ac
FT exon 8835..8943
```

```
FT intron /tag= ad
FT /number= "9"
FT 8944..9092
FT *tag= ae
FT /number= "9"
FT allele replace(8951,C)
FT *tag= af
FT exon 9093..9151
FT *tag= ag
FT /number= "10"
FT intron 9152..9239
FT *tag= ah
FT /number= "10"
FT allele replace(9154,G)
FT *tag= ai
FT exon 9240..9309
FT *tag= aj
FT /number= "11"
FT allele replace(9431,G)
FT *tag= ak
```

WO200202579-A2.

10-JAN-2002.

02-JUL-2001; 2001WO-US20952.

30-JUN-2000; 2000US-215236P.

(GENA-) GENAISSANCE PHARM INC.

Duda A, Finkel K, Koshy B, Parks KE;

WPI; 2002-154721/20.

P-FSDB; AAM48922.

Novel genetic variants of enolase 3, (beta, muscle) gene useful in studying expression and function of the protein, and for screening drugs to treat disorders of glycolytic pathway -

Claim 20; Fig 1; 90pp; English.

The present invention provides the protein, cDNA and genomic sequences of a human enolase 3 (beta, muscle) isogene containing a number of single nucleotide polymorphisms (SNPs). The sequences can be used to identify the haplotype of an individual and identify whether particular haplotypes are linked to certain diseases. The present sequence is the ENO3 gene.

SQ Sequence 13308 BP; 2874 A; 3775 C; 3758 G; 2901 T; 0 other;

Query Match 76.4%; Score 16.8; DB 24; Length 13308;

Best Local Similarity 90.0%; Pred. No. 3.9e+02;

Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 CCTTGGCGCGCTGCGCTCA 21

Db 1405 CCGTGGCGCGCTGCGCTCA 1424

Search completed: April 18, 2003, 05:45:23

Job time : 103 secs

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 18, 2003, 04:53:41 ; Search time 21.5 Seconds
(without alignments)
313.809 Million cell updates/sec

Title: US-09-270-437D-11

Perfect score: 22
Sequence: 1 tccttgcgcgtcgccctcag 22

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- Issued Patents NA:*
- 1: /cgn2_6/prodata/1/ina/5A COMB.seq:*
 - 2: /cgn2_6/prodata/1/ina/5B COMB.seq:*
 - 3: /cgn2_6/prodata/1/ina/6A COMB.seq:*
 - 4: /cgn2_6/prodata/1/ina/6B COMB.seq:*
 - 5: /cgn2_6/prodata/1/ina/PCUS COMB.seq:*
 - 6: /cgn2_6/prodata/1/ina/backfile1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	22	100.0	1946	4	US-09-061-709-7
2	17.2	78.2	1704	4	US-09-608-917A-3
3	16.8	76.4	462	4	US-09-222-575-165
4	16.2	73.6	2409	4	US-09-230-225B-3
5	15.8	71.8	1471	4	US-09-492-985-11
6	15.8	71.8	3744	2	US-08-348-353-16
7	15.8	71.8	3744	2	US-08-465-965-16
8	15.8	71.8	3744	2	US-08-465-965-16
9	15.6	70.9	783	3	US-08-264-861A-11
10	15.6	70.9	783	5	PCT-US95-07784-11
11	15.6	70.9	1150	1	US-08-264-861A-10
12	15.6	70.9	1150	5	PCT-US95-07784-10
13	15.4	70.0	1263	2	US-08-828-488-6
14	15.4	70.0	1263	2	US-09-299-689A-6
15	15.2	69.1	2936	2	US-08-738-172-1
16	15.2	69.1	423	1	US-08-470-179-127
17	15.2	69.1	1588	2	US-08-706-037-24
18	15.2	69.1	1588	2	US-09-005-397-24
19	15.2	69.1	1672	1	US-08-172-331B-13
20	15.2	69.1	1943	4	US-09-282-305-7
21	15.2	69.1	3175	1	US-08-050-684-1
22	15.2	69.1	3175	1	US-08-582-719-1
23	15.2	69.1	5919	4	US-09-221-017B-584
24	15.2	69.1	80161	3	US-09-036-987A-1
25	15.2	69.1	80161	4	US-09-370-700-1
26	14.8	67.3	391	3	US-08-516-859A-104
27	14.8	67.3	391	4	US-09-586-472-104

C 28	14.8	67.3	391	4	US-09-528-706-104	Sequence 104, Appl
C 29	14.8	67.3	600	4	US-09-452-239-39	Sequence 39, Appl
C 30	14.8	67.3	750	4	US-09-318-191-33	Sequence 33, Appl
C 31	14.8	67.3	1078	4	US-09-452-239-41	Sequence 41, Appl
C 32	14.8	67.3	1334	1	US-08-395-742-1	Sequence 1, Appl
C 33	14.8	67.3	2244	1	US-08-476-519-10	Sequence 10, Appl
C 34	14.8	67.3	2244	5	PCT-US95-09323-10	Sequence 10, Appl
C 35	14.8	67.3	2334	1	US-08-476-519-1	Sequence 1, Appl
C 36	14.8	67.3	2334	5	PCT-US95-09323-1	Sequence 1, Appl
C 37	14.8	67.3	2594	3	US-08-989-385-2	Sequence 2, Appl
C 38	14.8	67.3	4190	4	US-08-924-345-1	Sequence 2, Appl
C 39	14.8	67.3	403765	4	US-09-103-840A-2	Sequence 2, Appl
C 40	14.6	66.4	434	2	US-08-997-080-50	Sequence 50, Appl
C 41	14.6	66.4	434	2	US-08-997-362-50	Sequence 50, Appl
C 42	14.6	66.4	434	3	US-08-873-970-50	Sequence 50, Appl
C 43	14.6	66.4	434	4	US-09-095-855-50	Sequence 50, Appl
C 44	14.6	66.4	434	4	US-08-705-347A-50	Sequence 50, Appl
C 45	14.6	66.4	434	4	US-09-324-542-50	Sequence 50, Appl

ALIGNMENTS

RESULT 1

US-09-061-709-7
; Sequence 7, Application US/09061709B
; Patent No. 6297364

; GENERAL INFORMATION:

; APPLICANT: Chen, Yao-Tseng

; APPLICANT: Gure, Ali

; APPLICANT: Tsang, Solam

; APPLICANT: Stockert, Elisabeth

; APPLICANT: Jager, Elke

; APPLICANT: Knuth, Alexander

; APPLICANT: Old, Lloyd J.

; TITLE OF INVENTION: Isolated Nucleic Acid Molecules Encoding Cancer Associated

; FILE REFERENCE: LUD 5538

; CURRENT APPLICATION NUMBER: US/09/061,709B

; CURRENT FILING DATE: 1998-04-17

; NUMBER OF SEQ ID NOS: 8

; SEQ ID NO 7

; LENGTH: 1946

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

US-09-061-709-7

Query Match

Best Local Similarity 100.0%; Score 22; DB 4; Length 1946;

Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TCCTTGGCGCTCGCGCTCAG 22

Db 38 TCCTTGGCGCTCGCGCTCAG 59

RESULT 2

US-09-608-917A-3

; Sequence 3, Application US/09608917A

; Patent No. 6409648

; GENERAL INFORMATION:

; APPLICANT: The Regents of the University of California

; APPLICANT: Campisi, Judith

; APPLICANT: Kim, Sahn-Ho

; TITLE OF INVENTION: TRFI Binding Protein

; FILE REFERENCE: LBNL IB 1317

; CURRENT APPLICATION NUMBER: US/09/608,917A

; CURRENT FILING DATE: 2000-06-30

; NUMBER OF SEQ ID NOS: 4

; SOFTWARE: PatentIn version 3.0

; SEQ ID NO 3

; LENGTH: 1704

```
;
;
; TYPE: DNA
; ORGANISM: mouse
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1687)..(1687)
; OTHER INFORMATION: "n" represents "A" because it is in the poly A tail of the cDNA
US-09-608-917A-3

Query Match      78.2%; Score 17.2; DB 4; Length 1704;
Best Local Similarity 86.4%; Pred. No. 45;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 TCCTTGGCGGCTGCGGCTCAG 22
Db 319 TCCTTGGCGGCTGCTGCCCG 340

RESULT 3
US-09-222-575-165/c
; Sequence 165, Application US/09222575
; Patent No. 6387697
; GENERAL INFORMATION:
; APPLICANT: Yuguu, Jiang
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Xu, Jiangchun
; TITLE OF INVENTION: Compositions for the Treatment and Diagnosis of Breast Cancer
; TITLE OF INVENTION: and Methods for Their Use
; FILE REFERENCE: 210121.470
; CURRENT APPLICATION NUMBER: US/09/222,575
; CURRENT FILING DATE: 1998-12-28
; NUMBER OF SEQ ID NOS: 174
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 165
; LENGTH: 462
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (10)
; OTHER INFORMATION: Where n is a, c, g or t
; NAME/KEY: modified_base
; LOCATION: (33)
; OTHER INFORMATION: Where n is a, c, g or t
; NAME/KEY: modified_base
; LOCATION: (36)
; OTHER INFORMATION: Where n is a, c, g or t
; NAME/KEY: modified_base
; LOCATION: (49)
; OTHER INFORMATION: Where n is a, c, g or t
; NAME/KEY: modified_base
; LOCATION: (198)
; OTHER INFORMATION: Where n is a, c, g or t
; NAME/KEY: modified_base
; LOCATION: (222)
; OTHER INFORMATION: Where n is a, c, g or t
; NAME/KEY: modified_base
; LOCATION: (243)
; OTHER INFORMATION: Where n is a, c, g or t
; NAME/KEY: modified_base
; LOCATION: (278)
; OTHER INFORMATION: Where n is a, c, g or t
; NAME/KEY: modified_base
; LOCATION: (357)
; OTHER INFORMATION: Where n is a, c, g or t
; NAME/KEY: modified_base
; LOCATION: (385)
; OTHER INFORMATION: Where n is a, c, g or t
; NAME/KEY: modified_base
; LOCATION: (399)
; OTHER INFORMATION: Where n is a, c, g or t
; NAME/KEY: modified_base
; LOCATION: (405)

;
;
; OTHER INFORMATION: Where n is a, c, g or t
; NAME/KEY: modified_base
; LOCATION: (437)
; OTHER INFORMATION: Where n is a, c, g or t
US-09-222-575-165

Query Match      76.4%; Score 16.8; DB 4; Length 462;
Best Local Similarity 90.0%; Pred. No. 67;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 TCCTTGGCGGCTGCGGCTC 20
Db 91 TCCTTGGCTGCTGCGGCTC 72

RESULT 4
US-09-230-225B-3
; Sequence 3, Application US/09230225B
; Patent No. 6403362
; GENERAL INFORMATION:
; APPLICANT: Meiji Seika Kaisha, Ltd.
; APPLICANT: Moriya, Tatsuaki
; TITLE OF INVENTION: Systems for the Mass Production of Proteins or Peptides by Micro
; FILE REFERENCE: VX990054
; CURRENT APPLICATION NUMBER: US/09/230,225B
; CURRENT FILING DATE: 1999-03-03
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 2409
; TYPE: DNA
; ORGANISM: Humicola insolens
; FEATURE:
; NAME/KEY: sig peptide
; LOCATION: (389)..(457)
; OTHER INFORMATION:
; NAME/KEY: mat peptide
; LOCATION: (458)..()
; OTHER INFORMATION:
; NAME/KEY: CDS
; LOCATION: (458)..(477)
; OTHER INFORMATION:
; NAME/KEY: Intron
; LOCATION: (478)..(535)
; OTHER INFORMATION:
; NAME/KEY: CDS
; LOCATION: (536)..(1029)
; OTHER INFORMATION:
; NAME/KEY: Intron
; LOCATION: (1030)..(1141)
; OTHER INFORMATION:
; NAME/KEY: CDS
; LOCATION: (1142)..(1761)
; OTHER INFORMATION:
; NAME/KEY: Intron
; LOCATION: (1762)..(1815)
; OTHER INFORMATION:
; NAME/KEY: CDS
; LOCATION: (1816)..(1989)
; OTHER INFORMATION:
; NAME/KEY: Intron
; LOCATION: (1990)..(2044)
; OTHER INFORMATION:
; NAME/KEY: CDS
; LOCATION: (2045)..(2095)
; OTHER INFORMATION:
; NAME/KEY: misc feature
; LOCATION: (688)..(693)
; OTHER INFORMATION: Cleavage site SmaI
; NAME/KEY: misc feature
; LOCATION: (1253)..(1259)
; OTHER INFORMATION: Cleavage site BamHI
```

NAME/KEY: misc.feature
LOCATION: (1505)..(1510)
OTHER INFORMATION: Cleavage site BgIII
NAME/KEY: misc.feature
LOCATION: (1643)..(1648)
OTHER INFORMATION: Cleavage siteStuI
US-09-230-225B-3

Query Match 73.6%; Score 16.2; DB 4; Length 2409;
Best Local Similarity 85.7%; Pred. No. 1.2e+02;
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 2 CCTTGGCGCTCGCGCTCAG 22
Db 886 CCTTGGCGCTCGCGCTCGG 906

RESULT 5
US-09-492-985-11/c
Sequence 11 Application US/09492985
Patent No. 6376240
GENERAL INFORMATION:
APPLICANT: Song, An M.
APPLICANT: Chen, Ya-Pen
APPLICANT: Krensky, Alan M.
TITLE OF INVENTION: RPLAT-1: A Transcription Factor That
TITLE OF INVENTION: Activates RANTES Gene Expression
FILE REFERENCE: SUN-113P
CURRENT APPLICATION NUMBER: US/09/492,985
CURRENT FILING DATE: 2000-01-27
EARLIER APPLICATION NUMBER: 60/117,576
EARLIER FILING DATE: 1999-01-27
NUMBER OF SEQ ID NOS: 11
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 11
LENGTH: 1471
TYPE: DNA
ORGANISM: mouse
US-09-492-985-11

Query Match 71.8%; Score 15.8; DB 4; Length 1471;
Best Local Similarity 89.5%; Pred. No. 1.7e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 CCTTGGCGCTCGCGCTC 20
Db 650 CCTTGGCGCAGCGGCC 632

RESULT 6
US-08-348-353-16/c
Sequence 16 Application US/08348353
Patent No. 5932217
GENERAL INFORMATION:
APPLICANT: Tuomanen, Elaine
APPLICANT: Masure, Robert
TITLE OF INVENTION: Antibody Recognizing Endothelial Cell
TITLE OF INVENTION: Ligand for Leukocyte CR3
NUMBER OF SEQUENCES: 38
CORRESPONDENCE ADDRESS:
ADDRESSEE: Klauber & Jackson
STREET: 411 Hackensack Avenue
CITY: Hackensack
STATE: New Jersey
COUNTRY: U.S.A.
ZIP: 07601
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/348,353

FILING DATE: 30-NOV-1994
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Jackson, David
REGISTRATION NUMBER: 26,742
REFERENCE/DOCKET NUMBER: 600-1-097CIP1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201-487-5800
TELEFAX: 201-343-1684
INFORMATION FOR SEQ ID NO: 16:
SEQUENCE CHARACTERISTICS:
LENGTH: 3744 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
FEATURE:
NAME/KEY: CDS
LOCATION: 1..3744
PUBLICATION INFORMATION:
AUTHORS: Delisse-Gathoye, et al.
TITLE: Cloning, Partial Sequence, Expressions, and
TITLE: Antigenic Analysis of the Filamentous
TITLE: Hemagglutinin Gene of Bordetella Pertussis
JOURNAL: Infection and Immunity
VOLUME: 58
ISSUE: 9
PAGES: 2895-2905
DATE: September-1990
US-08-348-353-16

Query Match 71.8%; Score 15.8; DB 2; Length 3744;
Best Local Similarity 89.5%; Pred. No. 1.7e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 TCCTTGGCGCTCGCGCT 19
Db 3050 TCCTTGGCGCTCGGCAT 3032

RESULT 7
US-08-465-965-16/c
Sequence 16 Application US/08465965
Patent No. 5968512
GENERAL INFORMATION:
APPLICANT: Tuomanen, Elaine
APPLICANT: Masure, Robert
TITLE OF INVENTION: Antibody Recognizing Endothelial Cell
TITLE OF INVENTION: Ligand for Leukocyte CR3
NUMBER OF SEQUENCES: 38
CORRESPONDENCE ADDRESS:
ADDRESSEE: Klauber & Jackson
STREET: 411 Hackensack Avenue
CITY: Hackensack
STATE: New Jersey
COUNTRY: U.S.A.
ZIP: 07601
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/465,965
FILING DATE: 06-JUN-1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/348,353
FILING DATE: 30-NOV-1994
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/247,572
FILING DATE: 23-MAY-1994
CLASSIFICATION: 424

us-09-270-437d-11.rni

Mon Apr 21 10:29:58 2003

```

PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US92/03725
FILING DATE: 04-MAY-1992
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/695,613
FILING DATE: 03-MAY-1991
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Jackson, David
REGISTRATION NUMBER: 26,742
REFERENCE/DOCKET NUMBER: 600-1-097CIP1DIV2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201-487-5800
TELEFAX: 201-343-1684
INFORMATION FOR SEQ ID NO: 16:
SEQUENCE CHARACTERISTICS:
LENGTH: 3744 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
FEATURE:
NAME/KEY: CDS
LOCATION: 1..3744
PUBLICATION INFORMATION:
AUTHORS: Delisse-Gathoye, et al.
TITLE: Cloning, Partial Sequence, Expressions, and
TITLE: Antigenic Analysis of the Filamentous
TITLE: Hemagglutinin Gene of Bordetella Pertussis
JOURNAL: Infection and Immunity
VOLUME: 58
ISSUE: 9
PAGES: 2895-2905
DATE: September-1990
US-08-465-965-16

Query Match 71.8%; Score 15.8; DB 2; Length 3744;
Best Local Similarity 89.5%; Pred. No. 1.7e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 TCCTTGGCGCTCGGCCT 19
Db 3050 TCCTTGGCGCTCGGCAT 3032

RESULT 8
US-08-465-966-16/C
Sequence 16, Application US/08465966
Patent No. 6015560
GENERAL INFORMATION:
APPLICANT: Tuomanen, Elaine
APPLICANT: Masure, Robert
TITLE OF INVENTION: Antibody Recognizing Endothelial Cell
TITLE OF INVENTION: Ligand for Leukocyte CR3
NUMBER OF SEQUENCES: 38
CORRESPONDENCE ADDRESS:
ADDRESSEE: Klauber & Jackson
STREET: 411 Hackensack Avenue
CITY: Hackensack
STATE: New Jersey
COUNTRY: U.S.A.
ZIP: 07601
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/465,966
FILING DATE: 06-JUN-1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/348,353

```

```

FILING DATE: 30-NOV-1994
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/247,572
FILING DATE: 23-MAY-1994
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US92/03725
FILING DATE: 04-MAY-1992
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/695,613
FILING DATE: 03-MAY-1991
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Jackson, David
REGISTRATION NUMBER: 26,742
REFERENCE/DOCKET NUMBER: 600-1-097CIP1DIV2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201-487-5800
TELEFAX: 201-343-1684
INFORMATION FOR SEQ ID NO: 16:
SEQUENCE CHARACTERISTICS:
LENGTH: 3744 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
FEATURE:
NAME/KEY: CDS
LOCATION: 1..3744
PUBLICATION INFORMATION:
AUTHORS: Delisse-Gathoye, et al.
TITLE: Cloning, Partial Sequence, Expressions, and
TITLE: Antigenic Analysis of the Filamentous
TITLE: Hemagglutinin Gene of Bordetella Pertussis
JOURNAL: Infection and Immunity
VOLUME: 58
ISSUE: 9
PAGES: 2895-2905
DATE: September-1990
US-08-465-966-16

Query Match 71.8%; Score 15.8; DB 3; Length 3744;
Best Local Similarity 89.5%; Pred. No. 1.7e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 TCCTTGGCGCTCGGCCT 19
Db 3050 TCCTTGGCGCTCGGCAT 3032

RESULT 9
US-08-264-861A-11/c
Sequence 11, Application US/08264861A
Patent No. 5622866
GENERAL INFORMATION:
APPLICANT: MOTAMED, HAIDEB
APPLICANT: SHAFIEE, ALI
TITLE OF INVENTION: EXPRESSION CASSETTES USEFUL IN
TITLE OF INVENTION: CONSTRUCTION OF INTEGRATIVE AND REPLICATIVE EXPRESSION
TITLE OF INVENTION: VECTORS FOR STREPTOMYCES
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESS:
ADDRESSEE: J. ERIC THIES
STREET: P.O. BOX 2000, 126 E. LINCOLN AVENUE
CITY: RAHWAY
STATE: NJ
COUNTRY: USA
ZIP: 07065
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS

```

;; SOFTWARE: PatentIn Release #1.0, Version #1.30
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/264,861A
;; FILING DATE: 23-JUN-1994
;; CLASSIFICATION: 435
;; ATTORNEY/AGENT INFORMATION:
;; NAME: THIES, J. ERIC
;; REGISTRATION NUMBER: 35,382
;; REFERENCE/DOCKET NUMBER: 19132
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (908)594-3904
;; TELEFAX: (908) 594-4720
;; INFORMATION FOR SEQ ID NO: 11:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 783 base pairs
;; TYPE: nucleic acid
;; STRANDEDNESS: single
;; TOPOLOGY: unknown
;; MOLECULE TYPE: DNA (genomic)
;; HYPOTHETICAL: NO
;; ANTI-SENSE: NO
US-08-264-861A-11

Query Match 70.9%; Score 15.6; DB 1; Length 783;
Best Local Similarity 81.8%; Pred. No. 2.1e+02;
Matches 18; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 TCCTTGGCGGCTGGCGCTCAG 22
DB 401 TCCTTGGCGGCGCGGCTCG 380

RESULT 10

PCT-US95-07784-11/c
; Sequence 11, Application PC/TUS9507784
; GENERAL INFORMATION:
; APPLICANT: MOTAMED, HAIDEH
; TITLE OF INVENTION: EXPRESSION CASSETTES USEFUL IN
; TITLE OF INVENTION: CONSTRUCTION OF INTEGRATIVE AND REPLICATIVE EXPRESSION
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: J. ERIC THIES
; STREET: P.O. BOX 2000, 126 E. LINCOLN AVENUE
; CITY: RAHWAY
; STATE: NJ
; COUNTRY: USA
; ZIP: 07065
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; FILING DATE: PCT/US95/07784
; CLASSIFICATION:
; NAME: THIES, J. ERIC
; REGISTRATION NUMBER: 35,382
; REFERENCE/DOCKET NUMBER: 19132
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (908)594-3904
; TELEFAX: (908) 594-4720
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 783 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: unknown
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO

; ANTI-SENSE: NO
PCT-US95-07784-11

Query Match 70.9%; Score 15.6; DB 5; Length 783;
Best Local Similarity 81.8%; Pred. No. 2.1e+02;
Matches 18; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 TCCTTGGCGGCTGGCGCTCAG 22
DB 401 TCCTTGGCGGCGCGGCTCG 380

RESULT 11

US-08-264-861A-10/c
; Sequence 10, Application US/08264861A
; Patent No. 5622866
; GENERAL INFORMATION:
; APPLICANT: MOTAMED, HAIDEH
; APPLICANT: SHAFIEE, ALI
; TITLE OF INVENTION: EXPRESSION CASSETTES USEFUL IN
; TITLE OF INVENTION: CONSTRUCTION OF INTEGRATIVE AND REPLICATIVE EXPRESSION
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: J. ERIC THIES
; STREET: P.O. BOX 2000, 126 E. LINCOLN AVENUE
; CITY: RAHWAY
; STATE: NJ
; COUNTRY: USA
; ZIP: 07065
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/264,861A
; FILING DATE: 23-JUN-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: THIES, J. ERIC
; REGISTRATION NUMBER: 35,382
; REFERENCE/DOCKET NUMBER: 19132
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (908)594-3904
; TELEFAX: (908) 594-4720
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1150 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: unknown
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
US-08-264-861A-10

Query Match 70.9%; Score 15.6; DB 1; Length 1150;
Best Local Similarity 81.8%; Pred. No. 2.1e+02;
Matches 18; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 TCCTTGGCGGCTGGCGCTCAG 22
DB 462 TCCTTGGCGGCGCGGCTCG 441

RESULT 12

PCT-US95-07784-10/c
; Sequence 10, Application PC/TUS9507784
; GENERAL INFORMATION:
; APPLICANT: MOTAMED, HAIDEH
; APPLICANT: SHAFIEE, ALI
; TITLE OF INVENTION: EXPRESSION CASSETTES USEFUL IN

TITLE OF INVENTION: CONSTRUCTION OF INTEGRATIVE AND REPLICATIVE EXPRESSION
VECTORS FOR STREPTOMYCES

TITLE OF INVENTION: 11
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: J. ERIC THIES
STREET: P.O. BOX 2000, 126 E. LINCOLN AVENUE
CITY: RAHWAY
STATE: NJ
COUNTRY: USA
ZIP: 07065

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
CURRENT APPLICATION NUMBER: PCT/US95/07784
FILING DATE:
CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:
NAME: THIES, J. ERIC
REGISTRATION NUMBER: 35,382
REFERENCE/DOCKET NUMBER: 19132
TELECOMMUNICATION INFORMATION:
TELEPHONE: (908) 594-3904
TELEFAX: (908) 594-4720

INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 1150 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: unknown
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
PCT-US95-07784-10

Query Match 70.9%; Score 15.6; DB 5; Length 1150;
Best Local Similarity 81.8%; Pred. No. 2.1e+02;
Matches 18; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

1 TCCTTGGCGGCTGGCGGCTCAG 22
|||||
462 TCCTTGGCGGCGGCGGCTCGC 441

RESULT 13
US-08-828-488-6
; Sequence 6, Application US/08828488
; Patent No. 5925521
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; APPLICANT: Hawkins, Phillip R.
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Lal, Preeti
; APPLICANT: Goli, Surya K.
; TITLE OF INVENTION: NOVEL HUMAN SERINE
; TITLE OF INVENTION: CARBOXYPEPTIDASE
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/828,488

STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: COLNCR101
CLONE: 770469
US-09-299-689A-6

Query Match 70.0%; Score 15.4; DB 4; Length 1263;
Best Local Similarity 94.1%; Pred. No. 2.5e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TCCTTGGCGGCTGCGGC 17
Db 19 TCCTTGGCGGCGCGGC 35

RESULT 15

US-08-738-172-1
; Sequence 1, Application US/08738172
; Patent No. 5939257
; GENERAL INFORMATION:
; APPLICANT: Szaez, Joseph
; APPLICANT: Davis, Maria
; TITLE OF INVENTION: THERMOSTABLE ALKALINE PHOSPHATASES
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: 633 West Fifth Street
; STREET: Suite 4700
; CITY: Los Angeles
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90071-2066

COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; MEDIUM TYPE: storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: IBM P.C. DOS 5.0
; SOFTWARE: Word Perfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/738,172
; FILING DATE:
; CLASSIFICATION: 435

PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/005,965
; FILING DATE: October 27, 1995
; APPLICATION NUMBER: 08/465,003
; FILING DATE: June 5, 1995
; APPLICATION NUMBER: 08/575,354
; FILING DATE: December 20, 1995
; APPLICATION NUMBER: 08/240,158
; FILING DATE: May 10, 1994
; APPLICATION NUMBER: 08/229,329
; FILING DATE: April 18, 1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Warburg, Richard J.
; REGISTRATION NUMBER: 32,327
; REFERENCE/DOCKET NUMBER: 222/158
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (213) 489-1600
; TELEFAX: (213) 955-0440
; TELEX: 67-3510

INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2936 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear

US-08-738-172-1

Query Match 70.0%; Score 15.4; DB 2; Length 2936;
Best Local Similarity 94.1%; Pred. No. 2.5e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TCCTTGGCGGCTGCGGC 17
Db 1111 TCCTTGGCGGCTGCGGC 1127

Search completed: April 18, 2003, 07:34:29
Job time : 27.5 secs

GenCore version 5.1.4.p5 4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 18, 2003, 06:35:31 ; Search time 91.3333 Seconds
(without alignments)
241.975 Million cell updates/sec

Title: US-09-270-437D-11

Perfect score: 22

Sequence: 1 TCCTTGCCTGCTGGCCTCAG 22

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 639749 seqs, 502280978 residues

Total number of hits satisfying chosen parameters: 1279498

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications NA:*

- 1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
- 3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
- 4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
- 6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
- 8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
- 10: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
- 11: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
- 12: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
- 13: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
- 14: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	22	100.0	1946	10	US-09-899-651-7
2	16.8	76.4	386	10	US-09-960-352-11656
3	16.8	76.4	462	9	US-10-076-622-165
4	16.8	76.4	462	10	US-09-604-287A-165
5	16.8	76.4	462	10	US-09-339-338-165
6	16.8	76.4	462	12	US-10-007-805-165
7	16.4	74.5	1641	10	US-09-808-387-27
8	16.4	74.5	1759	9	US-09-992-598-180
9	16.4	74.5	1759	9	US-09-989-293A-180
10	16.4	74.5	1759	9	US-09-990-444-180
11	16.4	74.5	1759	9	US-09-989-735-180
12	16.4	74.5	1759	9	US-09-989-730-180
13	16.4	74.5	1759	9	US-09-930-436-180
14	16.4	74.5	1759	9	US-09-931-181-180
15	16.4	74.5	1759	9	US-09-993-687-180
16	16.4	74.5	1759	9	US-09-989-734-180
17	16.4	74.5	1759	9	US-09-997-653-180
18	16.4	74.5	1759	9	US-10-174-590-151
19	16.4	74.5	1759	9	US-10-176-758-151

20	16.4	74.5	1759	9	US-10-175-737-151	Sequence 151, App
21	16.4	74.5	1759	9	US-09-993-667-180	Sequence 180, App
22	16.4	74.5	1759	9	US-10-173-706-151	Sequence 151, App
23	16.4	74.5	1759	9	US-10-175-738-151	Sequence 151, App
24	16.4	74.5	1759	9	US-10-175-752-151	Sequence 151, App
25	16.4	74.5	1759	9	US-10-176-482-151	Sequence 151, App
26	16.4	74.5	1759	9	US-10-176-757-151	Sequence 151, App
27	16.4	74.5	1759	9	US-10-176-913-151	Sequence 151, App
28	16.4	74.5	1759	9	US-10-180-552-151	Sequence 151, App
29	16.4	74.5	1759	9	US-10-180-557-151	Sequence 151, App
30	16.4	74.5	1759	9	US-09-990-438-180	Sequence 180, App
31	16.4	74.5	1759	9	US-09-990-562-180	Sequence 180, App
32	16.4	74.5	1759	9	US-09-997-428-180	Sequence 180, App
33	16.4	74.5	1759	9	US-09-997-666-180	Sequence 180, App
34	16.4	74.5	1759	9	US-10-173-700-151	Sequence 151, App
35	16.4	74.5	1759	9	US-10-174-572-151	Sequence 151, App
36	16.4	74.5	1759	9	US-10-174-579-151	Sequence 151, App
37	16.4	74.5	1759	9	US-10-174-582-151	Sequence 151, App
38	16.4	74.5	1759	9	US-10-174-588-151	Sequence 151, App
39	16.4	74.5	1759	9	US-10-175-739-151	Sequence 151, App
40	16.4	74.5	1759	9	US-10-175-740-151	Sequence 151, App
41	16.4	74.5	1759	9	US-10-175-743-151	Sequence 151, App
42	16.4	74.5	1759	9	US-10-176-488-151	Sequence 151, App
43	16.4	74.5	1759	9	US-10-176-492-151	Sequence 151, App
44	16.4	74.5	1759	9	US-10-176-747-151	Sequence 151, App
45	16.4	74.5	1759	9	US-10-176-750-151	Sequence 151, App

ALIGNMENTS

RESULT 1

US-09-899-651-7
; Sequence 7, Application US/09899651
; Patent No. US2002011470A1

; GENERAL INFORMATION:

; APPLICANT: Chen, Yao-Tseng

; APPLICANT: Teang, Solam

; APPLICANT: Stockert, Elisabeth

; APPLICANT: Jager, Elke

; APPLICANT: Knuth, Alexander

; APPLICANT: Old, Lloyd J.

; TITLE OF INVENTION: Isolated Nucleic Acid Molecules Encoding Cancer

; TITLE OF INVENTION: Associated

; TITLE OF INVENTION: Antigen, The Antigens Per Se, And Uses Thereof

; FILE REFERENCE: LUD 538

; CURRENT APPLICATION NUMBER: US/09/899,651

; PRIOR FILING DATE: 2001-07-06

; PRIOR APPLICATION NUMBER: US/09/061,709

; PRIOR FILING DATE: 1998-04-17

; NUMBER OF SEQ ID NOS: 8

; SEQ ID NO 7

; LENGTH: 1946

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

US-09-899-651-7

Query Match
Best Local Similarity 100.0%; Score 22; DB 10; Length 1946;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCCTTGCCTGCTGGCCTCAG 22

Db 38 TCCTTGCCTGCTGGCCTCAG 59

RESULT 2

US-09-960-352-11656/c

; Sequence 11656, Application US/09960352

; Patent No. US20020137139A1

; GENERAL INFORMATION:

us-09-270-437d-11.rnpb

Mon Apr 21 10:30:00 2003

```

; APPLICANT: Warren, Wesley C.
; APPLICANT: Tao, Nengbing
; APPLICANT: Byatt, John C.
; APPLICANT: Mathialagan, Nagappan
; TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
; TITLE OF INVENTION: MUSCLE AND FAT DEPOSITION
; FILE REFERENCE: 16511.006/37-21(10298)C
; CURRENT APPLICATION NUMBER: US/09/960,352
; CURRENT FILING DATE: 2001-09-24
; NUMBER OF SEQ ID NOS: 15112
; SEQ ID NO 11656
; LENGTH: 386
; TYPE: DNA
; ORGANISM: Bos taurus
; OTHER INFORMATION: Clone ID: 50-LIB3057-003-Q1-K1-E10
US-09-960-352-11656

Query Match          76.4%; Score 16.8; DB 10; Length 386;
Best Local Similarity 90.0%; Pred. No. 64;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 CCTTGGCGGCTGGCGCCTCA 21
    ||||| ||||| ||||| |||||
Db 326 CCTTGGCGGCTGGCGCCTCA 307

RESULT 3
US-10-076-622-165/c
; Sequence 165, Application US/10076622
; Publication No. US20030023036A1
; GENERAL INFORMATION:
; APPLICANT: Houghton, Raymond L.
; APPLICANT: Sleath, Paul R.
; APPLICANT: Persing, David H.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF BREAST CANCER
; FILE REFERENCE: 210121.470C11
; CURRENT APPLICATION NUMBER: US/10/076,622
; CURRENT FILING DATE: 2002-02-13
; NUMBER OF SEQ ID NOS: 627
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 165
; LENGTH: 462
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 10..33, 36..49, 198, 223, 243, 278, 357, 385, 399, 405, 437
; OTHER INFORMATION: n = A,T,C or G
US-10-076-622-165

Query Match          76.4%; Score 16.8; DB 9; Length 462;
Best Local Similarity 90.0%; Pred. No. 62;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 TCCTTGGCGGCTGGCGCCTC 20
    ||||| ||||| ||||| |||||
Db 91 TCTTTGCTGCTGGCGCCTC 72

RESULT 4
US-09-604-287A-165/c
; Sequence 165, Application US/09604287A
; Patent No. US20020064872A1
; GENERAL INFORMATION:
; APPLICANT: Jiang, Yuqiu
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Hepler, William T.
; APPLICANT: Henderson, Robert A.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Vedvick, Thomas S.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF BREAST CANCER

```

```

; FILE REFERENCE: 210121.470C7
; CURRENT APPLICATION NUMBER: US/09/604,287A
; CURRENT FILING DATE: 2000-06-22
; NUMBER OF SEQ ID NOS: 489
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 165
; LENGTH: 462
; TYPE: DNA
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(462)
; OTHER INFORMATION: n = A,T,C or G
US-09-604-287A-165

Query Match          76.4%; Score 16.8; DB 10; Length 462;
Best Local Similarity 90.0%; Pred. No. 62;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 TCCTTGGCGGCTGGCGCCTC 20
    ||||| ||||| ||||| |||||
Db 91 TCTTTGCTGCTGGCGCCTC 72

RESULT 5
US-09-339-338-165/c
; Sequence 165, Application US/09339338A
; Patent No. US20020102602A1
; GENERAL INFORMATION:
; APPLICANT: Yuqiu, Jiang
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Hepler, William T.
; APPLICANT: Henderson, Robert A.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Vedvick, Thomas S.
; TITLE OF INVENTION: COMPOSITIONS FOR THE TREATMENT AND
; TITLE OF INVENTION: DIAGNOSIS OF BREAST CANCER AND METHODS FOR THEIR USE
; FILE REFERENCE: 210121.470C2
; CURRENT APPLICATION NUMBER: US/09/339,338A
; CURRENT FILING DATE: 1999-06-23
; NUMBER OF SEQ ID NOS: 315
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 165
; LENGTH: 462
; TYPE: DNA
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(462)
; OTHER INFORMATION: n = A,T,C or G
US-09-339-338-165

Query Match          76.4%; Score 16.8; DB 10; Length 462;
Best Local Similarity 90.0%; Pred. No. 62;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 TCCTTGGCGGCTGGCGCCTC 20
    ||||| ||||| ||||| |||||
Db 91 TCTTTGCTGCTGGCGCCTC 72

RESULT 6
US-10-007-805-165/c
; Sequence 165, Application US/10007805
; Patent No. US20020150581A1
; GENERAL INFORMATION:
; APPLICANT: Jiang, Yuqiu
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Hepler, William T.
; APPLICANT: Henderson, Robert A.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Vedvick, Thomas S.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF BREAST CANCER

```

```

; APPLICANT: McNeill, Patricia D.
; APPLICANT: Durham, Margarita
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF BREAST CANCER
; FILE REFERENCE: 210121.470C10
; CURRENT APPLICATION NUMBER: US/10/007,805
; CURRENT FILING DATE: 2001-12-07
; NUMBER OF SEQ ID NOS: 593
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 165
; LENGTH: 462
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 10, 33, 36, 49, 198, 222, 243, 278, 357, 385, 399, 405, 437
; OTHER INFORMATION: n = A,T,C or G
US-10-007-805-165

```

```

Query Match          76.4%; Score 16.8; DB 12; Length 462;
Best Local Similarity 90.0%; Pred. No. 62;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 TCCTGGCGGCTGGCGCTC 20
Db 91 TCCTGGCGGCTGGCGCTC 72

```

```

RESULT 7
US-09-808-387-27
; Sequence 27, Application US/09808387
; Patent No. US20020132293A1
; GENERAL INFORMATION:
; APPLICANT: Kaia Palm
; APPLICANT: Tonis Timmusk
; APPLICANT: Cemines Research
; TITLE OF INVENTION: MAMMALIAN NEUTRALIZED FAMILY OF
; FILE REFERENCE: TRANSCRIPTION REGULATORS AND USES THEREFOR
; CURRENT APPLICATION NUMBER: US/09/808,387
; CURRENT FILING DATE: 2001-03-14
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 27
; LENGTH: 1641
; TYPE: DNA
; ORGANISM: Rat
US-09-808-387-27

```

```

Query Match          74.5%; Score 15.4; DB 10; Length 1641;
Best Local Similarity 94.4%; Pred. No. 74;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 TCCTTGGCGGCTGGCGGCC 18
Db 1011 TCCTGGCGGCTGGCGGCC 1028

```

```

RESULT 8
US-09-992-598-180
; Sequence 180, Application US/09992598
; Patent No. US20020160384A1
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi J.
; APPLICANT: Baker, Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnovers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Fong, Sherman
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey

```

```

; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Napier, Mary A.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tamas, Daniel
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE REFERENCE: P2730P1C20
; CURRENT APPLICATION NUMBER: US/09/992,598
; CURRENT FILING DATE: 2001-11-14
; PRIOR APPLICATION NUMBER: 60/049787
; PRIOR FILING DATE: 1997-06-16
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/065186
; PRIOR FILING DATE: 1997-11-12
; PRIOR APPLICATION NUMBER: 60/065311
; PRIOR FILING DATE: 1997-11-13
; PRIOR APPLICATION NUMBER: 60/066770
; PRIOR FILING DATE: 1997-11-24
; PRIOR APPLICATION NUMBER: 60/075945
; PRIOR FILING DATE: 1998-02-25
; PRIOR APPLICATION NUMBER: 60/078910
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/083322
; PRIOR FILING DATE: 1998-04-28
; PRIOR APPLICATION NUMBER: 60/084600
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/087106
; PRIOR FILING DATE: 1998-05-28
; PRIOR APPLICATION NUMBER: 60/087607
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087609
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087759
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087827
; PRIOR FILING DATE: 1998-06-03
; PRIOR APPLICATION NUMBER: 60/088021
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088025
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088026
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088028
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088029
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088030
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088033
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088326
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088167
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088202
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088212
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088217
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088655
; PRIOR FILING DATE: 1998-06-09

```

7	PRIOR FILING DATE: 1998-06-24	
7	PRIOR APPLICATION NUMBER: 60/090540	
7	PRIOR FILING DATE: 1998-06-24	
7	PRIOR APPLICATION NUMBER: 60/090542	
7	PRIOR FILING DATE: 1998-06-24	
7	PRIOR APPLICATION NUMBER: 60/090557	
7	PRIOR FILING DATE: 1998-06-24	
7	PRIOR APPLICATION NUMBER: 60/090676	
7	PRIOR FILING DATE: 1998-06-25	
7	PRIOR APPLICATION NUMBER: 60/090678	
7	PRIOR FILING DATE: 1998-06-25	
7	PRIOR APPLICATION NUMBER: 60/090690	
7	PRIOR FILING DATE: 1998-06-25	
7	PRIOR APPLICATION NUMBER: 60/090694	
7	PRIOR FILING DATE: 1998-06-25	
7	PRIOR APPLICATION NUMBER: 60/090695	
7	PRIOR FILING DATE: 1998-06-25	
7	PRIOR APPLICATION NUMBER: 60/090696	
7	PRIOR FILING DATE: 1998-06-25	
7	PRIOR APPLICATION NUMBER: 60/090862	
7	PRIOR FILING DATE: 1998-06-26	
7	PRIOR APPLICATION NUMBER: 60/090863	
7	PRIOR FILING DATE: 1998-06-26	
7	PRIOR APPLICATION NUMBER: 60/091360	
7	PRIOR FILING DATE: 1998-07-01	
7	PRIOR APPLICATION NUMBER: 60/091478	
7	PRIOR FILING DATE: 1998-07-02	
7	PRIOR APPLICATION NUMBER: 60/091544	
7	PRIOR FILING DATE: 1998-07-01	
7	PRIOR APPLICATION NUMBER: 60/091519	
7	PRIOR FILING DATE: 1998-07-02	
7	PRIOR APPLICATION NUMBER: 60/091626	
7	PRIOR FILING DATE: 1998-07-02	
7	PRIOR APPLICATION NUMBER: 60/091633	
7	PRIOR FILING DATE: 1998-07-02	
7	PRIOR APPLICATION NUMBER: 60/091787	
7	PRIOR FILING DATE: 1998-07-07	
7	PRIOR APPLICATION NUMBER: 60/091822	
7	PRIOR FILING DATE: 1998-07-09	

Qy 5 TCGCGGCTGCGGCTCAG 22
nb 98 TCGCGGCTGCTGCCTCAG 115

```

RESULT 9
US-09-989-293A-180
; Sequence 180, Application US/09989293A
; Patent No. US20020177164A1
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi J.
; APPLICANT: Baker, Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Fong, Sherman
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Kijavini, Ivar J.
; APPLICANT: Napier, Mary A.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.

```

APPLICANT: Roy, Margaret Ann
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Watanabe, Colin K.
APPLICANT: Williams, P. Mickey
APPLICANT: Wood, William I.
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
FILE OF INVENTION: Acids Encoding the Same
FILE REFERENCE: P2730P1C66
CURRENT APPLICATION NUMBER: US/09/989,293A
CURRENT FILING DATE: 2001-11-20
PRIOR APPLICATION NUMBER: 60/049787
PRIOR FILING DATE: 1997-06-16
PRIOR APPLICATION NUMBER: 60/062250
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/065186
PRIOR FILING DATE: 1997-11-12
PRIOR APPLICATION NUMBER: 60/065311
PRIOR FILING DATE: 1997-11-13
PRIOR APPLICATION NUMBER: 60/066770
PRIOR FILING DATE: 1997-11-24
PRIOR APPLICATION NUMBER: 60/075945
PRIOR FILING DATE: 1998-02-25
PRIOR APPLICATION NUMBER: 60/078910
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/083322
PRIOR FILING DATE: 1998-04-28
PRIOR APPLICATION NUMBER: 60/084600
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/087106
PRIOR FILING DATE: 1998-05-28
PRIOR APPLICATION NUMBER: 60/087607
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087609
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087759
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087827
PRIOR FILING DATE: 1998-06-03
PRIOR APPLICATION NUMBER: 60/088021
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088025
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088026
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088028
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088029
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088030
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088033
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088326
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088167
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088202
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088212
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088217
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088655
PRIOR FILING DATE: 1998-06-09
PRIOR APPLICATION NUMBER: 60/088734
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088738
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088742
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088810
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088824
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088826
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088826
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088858
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/088861
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/088876
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/089105
PRIOR FILING DATE: 1998-06-12
PRIOR APPLICATION NUMBER: 60/089440
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089512
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089514
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089532
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089538
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089598
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089599
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089600
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089653
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089801
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089907
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089908
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089947
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/089948
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/089952
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/090246
PRIOR FILING DATE: 1998-06-22
PRIOR APPLICATION NUMBER: 60/090252
PRIOR FILING DATE: 1998-06-22
PRIOR APPLICATION NUMBER: 60/090254
PRIOR FILING DATE: 1998-06-22
PRIOR APPLICATION NUMBER: 60/090349
PRIOR FILING DATE: 1998-06-23
PRIOR APPLICATION NUMBER: 60/090355
PRIOR FILING DATE: 1998-06-23
PRIOR APPLICATION NUMBER: 60/090429
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090431
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090435
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090444
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090445
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090472
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090535
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090540
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090542
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090557
PRIOR FILING DATE: 1998-06-24

us-09-270-437d-11.rnpb

Mon Apr 21 10:30:00 2003

```

; PRIOR APPLICATION NUMBER: 60/090676
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090678
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090690
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090694
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090695
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090696
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090862
; PRIOR FILING DATE: 1998-06-26
; PRIOR APPLICATION NUMBER: 60/090863
; PRIOR FILING DATE: 1998-06-26
; PRIOR APPLICATION NUMBER: 60/091360
; PRIOR FILING DATE: 1998-07-01
; PRIOR APPLICATION NUMBER: 60/091478
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091544
; PRIOR FILING DATE: 1998-07-01
; PRIOR APPLICATION NUMBER: 60/091519
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091626
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091633
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091978
; PRIOR FILING DATE: 1998-07-07
; PRIOR APPLICATION NUMBER: 60/091982
; PRIOR FILING DATE: 1998-07-07
; PRIOR APPLICATION NUMBER: 60/092182
; PRIOR FILING DATE: 1998-07-09

Query Match          74.5%; Score 16.4; DB 9; Length 1759;
Best Local Similarity 94.4%; Pred No. 73;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 5 TGCGCGCTGCGGCTCAG 22
   ||||| ||||| |||||
DB 98 TGCGCGCTGCTGCTCAG 115

RESULT 10
US-09-989-735-180
; Sequence 180, Application US/09989735
; Publication No. US20020193299A1
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi J.
; APPLICANT: Baker, Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Fong, Sherman
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Kijavini, Ivar J.
; APPLICANT: Napier, Mary A.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tunas, Daniel
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE REFERENCE: P2730PIC61
; CURRENT APPLICATION NUMBER: US/09/989,735
; PRIOR FILING DATE: 2001-11-19
; PRIOR APPLICATION NUMBER: 60/049787
; PRIOR FILING DATE: 1997-06-16
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/065186
; PRIOR FILING DATE: 1997-11-12
; PRIOR APPLICATION NUMBER: 60/065311
; PRIOR FILING DATE: 1997-11-13
; PRIOR APPLICATION NUMBER: 60/066770
; PRIOR FILING DATE: 1997-11-24
; PRIOR APPLICATION NUMBER: 60/075945
; PRIOR FILING DATE: 1998-02-25
; PRIOR APPLICATION NUMBER: 60/078910
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/083322
; PRIOR FILING DATE: 1998-04-28
; PRIOR APPLICATION NUMBER: 60/084600
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/087106
; PRIOR FILING DATE: 1998-05-28
; PRIOR APPLICATION NUMBER: 60/087607
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087609
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087759
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087827
; PRIOR FILING DATE: 1998-06-03
; PRIOR APPLICATION NUMBER: 60/088021
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088025
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088026
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088028
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088029
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088030
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088033
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088326
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088167
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088202
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088212
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088217
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088655
; PRIOR FILING DATE: 1998-06-09
; PRIOR APPLICATION NUMBER: 60/088734
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088738
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088742
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088810
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088824
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088826
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088858
; PRIOR FILING DATE: 1998-06-11
```


;; PRIOR APPLICATION NUMBER: 60/088861
;; PRIOR FILING DATE: 1998-06-11
;; PRIOR APPLICATION NUMBER: 60/088876
;; PRIOR FILING DATE: 1998-06-11
;; PRIOR APPLICATION NUMBER: 60/089105
;; PRIOR FILING DATE: 1998-06-12
;; PRIOR APPLICATION NUMBER: 60/089440
;; PRIOR FILING DATE: 1998-06-16
;; PRIOR APPLICATION NUMBER: 60/089512
;; PRIOR FILING DATE: 1998-06-16
;; PRIOR APPLICATION NUMBER: 60/089514
;; PRIOR FILING DATE: 1998-06-16
;; PRIOR APPLICATION NUMBER: 60/089532
;; PRIOR FILING DATE: 1998-06-17
;; PRIOR APPLICATION NUMBER: 60/089538
;; PRIOR FILING DATE: 1998-06-17
;; PRIOR APPLICATION NUMBER: 60/089598
;; PRIOR FILING DATE: 1998-06-17
;; PRIOR APPLICATION NUMBER: 60/089599
;; PRIOR FILING DATE: 1998-06-17
;; PRIOR APPLICATION NUMBER: 60/089600
;; PRIOR FILING DATE: 1998-06-17
;; PRIOR APPLICATION NUMBER: 60/089653
;; PRIOR FILING DATE: 1998-06-17
;; PRIOR APPLICATION NUMBER: 60/089801
;; PRIOR FILING DATE: 1998-06-18
;; PRIOR APPLICATION NUMBER: 60/089907
;; PRIOR FILING DATE: 1998-06-18
;; PRIOR APPLICATION NUMBER: 60/089908
;; PRIOR FILING DATE: 1998-06-18
;; PRIOR APPLICATION NUMBER: 60/089947
;; PRIOR FILING DATE: 1998-06-19
;; PRIOR APPLICATION NUMBER: 60/089948
;; PRIOR FILING DATE: 1998-06-19
;; PRIOR APPLICATION NUMBER: 60/089952
;; PRIOR FILING DATE: 1998-06-19
;; PRIOR APPLICATION NUMBER: 60/090246
;; PRIOR FILING DATE: 1998-06-22
;; PRIOR APPLICATION NUMBER: 60/090252
;; PRIOR FILING DATE: 1998-06-22
;; PRIOR APPLICATION NUMBER: 60/090254
;; PRIOR FILING DATE: 1998-06-22
;; PRIOR APPLICATION NUMBER: 60/090349
;; PRIOR FILING DATE: 1998-06-23
;; PRIOR APPLICATION NUMBER: 60/090355
;; PRIOR FILING DATE: 1998-06-23
;; PRIOR APPLICATION NUMBER: 60/090429
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090431
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090435
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090444
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090445
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090472
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090535
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090540
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090542
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090557
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090676
;; PRIOR FILING DATE: 1998-06-25
;; PRIOR APPLICATION NUMBER: 60/090678
;; PRIOR FILING DATE: 1998-06-25
;; PRIOR APPLICATION NUMBER: 60/090690
;; PRIOR FILING DATE: 1998-06-25
;; PRIOR APPLICATION NUMBER: 60/090694

;; PRIOR FILING DATE: 1998-06-25
;; PRIOR APPLICATION NUMBER: 60/090695
;; PRIOR FILING DATE: 1998-06-25
;; PRIOR APPLICATION NUMBER: 60/090696
;; PRIOR FILING DATE: 1998-06-25
;; PRIOR APPLICATION NUMBER: 60/090862
;; PRIOR FILING DATE: 1998-06-26
;; PRIOR APPLICATION NUMBER: 60/090863
;; PRIOR FILING DATE: 1998-06-26
;; PRIOR APPLICATION NUMBER: 60/091360
;; PRIOR FILING DATE: 1998-07-01
;; PRIOR APPLICATION NUMBER: 60/091478
;; PRIOR FILING DATE: 1998-07-02
;; PRIOR APPLICATION NUMBER: 60/091544
;; PRIOR FILING DATE: 1998-07-01
;; PRIOR APPLICATION NUMBER: 60/091519
;; PRIOR FILING DATE: 1998-07-02
;; PRIOR APPLICATION NUMBER: 60/091626
;; PRIOR FILING DATE: 1998-07-02
;; PRIOR APPLICATION NUMBER: 60/091633
;; PRIOR FILING DATE: 1998-07-02
;; PRIOR APPLICATION NUMBER: 60/09178
;; PRIOR FILING DATE: 1998-07-07
;; PRIOR APPLICATION NUMBER: 60/091982
;; PRIOR FILING DATE: 1998-07-07
;; PRIOR APPLICATION NUMBER: 60/092182
;; PRIOR FILING DATE: 1998-07-09

Query Match 74.5%; Score 16.4; DB 9; Length 1759;
Best Local Similarity 94.4%; Pred. No. 73;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 5 TCGCGCTGCGGCTCAG 22
Dd 98 TCGCGCTGCTGCTCAG 115
|||||||

RESULT 11
US-09-990-444-180
; Sequence 180, Application US/09990444
; Publication No. US20020193300A1
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi J.
; APPLICANT: Baker, Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Fong, Sherman
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Napier, Mary A.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas P.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; TITLE OF INVENTION: Acids Encoding the Same
; FILE REFERENCE: E2730PIC19
; CURRENT APPLICATION NUMBER: US/09/990,444
; CURRENT FILING DATE: 2001-11-14
; PRIOR APPLICATION NUMBER: 60/049787
; PRIOR FILING DATE: 1997-06-16

us-09-270-437d-11.rnpb

Mon Apr 21 10:30:00 2003

, PRIOR APPLICATION NUMBER: 60/062250
 , PRIOR FILING DATE: 1997-10-17
 , PRIOR APPLICATION NUMBER: 60/065186
 , PRIOR FILING DATE: 1997-11-12
 , PRIOR APPLICATION NUMBER: 60/065311
 , PRIOR FILING DATE: 1997-11-13
 , PRIOR APPLICATION NUMBER: 60/066770
 , PRIOR FILING DATE: 1997-11-24
 , PRIOR APPLICATION NUMBER: 60/075945
 , PRIOR FILING DATE: 1998-02-25
 , PRIOR APPLICATION NUMBER: 60/078910
 , PRIOR FILING DATE: 1998-03-20
 , PRIOR APPLICATION NUMBER: 60/083322
 , PRIOR FILING DATE: 1998-04-28
 , PRIOR APPLICATION NUMBER: 60/084600
 , PRIOR FILING DATE: 1998-05-07
 , PRIOR APPLICATION NUMBER: 60/087106
 , PRIOR FILING DATE: 1998-05-28
 , PRIOR APPLICATION NUMBER: 60/087607
 , PRIOR FILING DATE: 1998-06-02
 , PRIOR APPLICATION NUMBER: 60/087609
 , PRIOR FILING DATE: 1998-06-02
 , PRIOR APPLICATION NUMBER: 60/087759
 , PRIOR FILING DATE: 1998-06-02
 , PRIOR APPLICATION NUMBER: 60/087827
 , PRIOR FILING DATE: 1998-06-03
 , PRIOR APPLICATION NUMBER: 60/088021
 , PRIOR FILING DATE: 1998-06-04
 , PRIOR APPLICATION NUMBER: 60/088025
 , PRIOR FILING DATE: 1998-06-04
 , PRIOR APPLICATION NUMBER: 60/088026
 , PRIOR FILING DATE: 1998-06-04
 , PRIOR APPLICATION NUMBER: 60/088028
 , PRIOR FILING DATE: 1998-06-04
 , PRIOR APPLICATION NUMBER: 60/088029
 , PRIOR FILING DATE: 1998-06-04
 , PRIOR APPLICATION NUMBER: 60/088030
 , PRIOR FILING DATE: 1998-06-04
 , PRIOR APPLICATION NUMBER: 60/088033
 , PRIOR FILING DATE: 1998-06-04
 , PRIOR APPLICATION NUMBER: 60/088326
 , PRIOR FILING DATE: 1998-06-04
 , PRIOR APPLICATION NUMBER: 60/088167
 , PRIOR FILING DATE: 1998-06-05
 , PRIOR APPLICATION NUMBER: 60/088202
 , PRIOR FILING DATE: 1998-06-05
 , PRIOR APPLICATION NUMBER: 60/088212
 , PRIOR FILING DATE: 1998-06-05
 , PRIOR APPLICATION NUMBER: 60/088217
 , PRIOR FILING DATE: 1998-06-05
 , PRIOR APPLICATION NUMBER: 60/088655
 , PRIOR FILING DATE: 1998-06-09
 , PRIOR APPLICATION NUMBER: 60/088734
 , PRIOR FILING DATE: 1998-06-10
 , PRIOR APPLICATION NUMBER: 60/088738
 , PRIOR FILING DATE: 1998-06-10
 , PRIOR APPLICATION NUMBER: 60/088742
 , PRIOR FILING DATE: 1998-06-10
 , PRIOR APPLICATION NUMBER: 60/088810
 , PRIOR FILING DATE: 1998-06-10
 , PRIOR APPLICATION NUMBER: 60/088824
 , PRIOR FILING DATE: 1998-06-10
 , PRIOR APPLICATION NUMBER: 60/088826
 , PRIOR FILING DATE: 1998-06-10
 , PRIOR APPLICATION NUMBER: 60/088858
 , PRIOR FILING DATE: 1998-06-11
 , PRIOR APPLICATION NUMBER: 60/088861
 , PRIOR FILING DATE: 1998-06-11
 , PRIOR APPLICATION NUMBER: 60/088876
 , PRIOR FILING DATE: 1998-06-11
 , PRIOR APPLICATION NUMBER: 60/089105
 , PRIOR FILING DATE: 1998-06-12
 , PRIOR APPLICATION NUMBER: 60/089440

; PRIOR APPLICATION NUMBER: 60/090863
 ; PRIOR FILING DATE: 1998-06-26
 ; PRIOR APPLICATION NUMBER: 60/091360
 ; PRIOR FILING DATE: 1998-07-01
 ; PRIOR APPLICATION NUMBER: 60/091478
 ; PRIOR FILING DATE: 1998-07-02
 ; PRIOR APPLICATION NUMBER: 60/091544
 ; PRIOR FILING DATE: 1998-07-01
 ; PRIOR APPLICATION NUMBER: 60/091519
 ; PRIOR FILING DATE: 1998-07-02
 ; PRIOR APPLICATION NUMBER: 60/091626
 ; PRIOR FILING DATE: 1998-07-02
 ; PRIOR APPLICATION NUMBER: 60/091633
 ; PRIOR FILING DATE: 1998-07-02
 ; PRIOR APPLICATION NUMBER: 60/091978
 ; PRIOR FILING DATE: 1998-07-07
 ; PRIOR APPLICATION NUMBER: 60/091982
 ; PRIOR FILING DATE: 1998-07-07
 ; PRIOR APPLICATION NUMBER: 60/092182
 ; PRIOR FILING DATE: 1998-07-09

Query Match 74.5%; Score 16.4; DB 9; Length 1759;
 Best Local Similarity 94.4%; Pred. No. 73;
 Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 5 TCGCGCTGCGGCTCTCAG 22
 Db 98 TCGCGCTGCGGCTCTCAG 115

RESULT 12

US-09-989-730-180
 ; Sequence 180, Application US/09989730
 ; Publication No US200020197674A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Ashkenazi, Avi J.
 ; APPLICANT: Baker, Kevin P.
 ; APPLICANT: Botstein, David
 ; APPLICANT: Desnoyers, Luc
 ; APPLICANT: Eaton, Dan L.
 ; APPLICANT: Ferrara, Napoleone
 ; APPLICANT: Fong, Sherman
 ; APPLICANT: Gerber, Hanspeter
 ; APPLICANT: Gerritsen, Mary E.
 ; APPLICANT: Goddard, Audrey
 ; APPLICANT: Godowski, Paul J.
 ; APPLICANT: Grimaldi, J. Christopher
 ; APPLICANT: Gurney, Austin L.
 ; APPLICANT: Kljavin, Ivar J.
 ; APPLICANT: Napier, Mary A.
 ; APPLICANT: Pan, James
 ; APPLICANT: Paoni, Nicholas F.
 ; APPLICANT: Roy, Margaret Ann
 ; APPLICANT: Stewart, Timothy A.
 ; APPLICANT: Tamas, Daniel
 ; APPLICANT: Watanabe, Colin K.
 ; APPLICANT: Williams, P. Mickey
 ; APPLICANT: Wood, William I.
 ; APPLICANT: Zhang, Zemin
 ; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
 ; FILE REFERENCE: P27301C9
 ; CURRENT APPLICATION NUMBER: US/09/989,730
 ; PRIOR FILING DATE: 2001-11-20
 ; PRIOR APPLICATION NUMBER: 60/049787
 ; PRIOR FILING DATE: 1997-06-16
 ; PRIOR APPLICATION NUMBER: 60/062250
 ; PRIOR FILING DATE: 1997-10-17
 ; PRIOR APPLICATION NUMBER: 60/055186
 ; PRIOR FILING DATE: 1997-11-12
 ; PRIOR APPLICATION NUMBER: 60/065311
 ; PRIOR FILING DATE: 1997-11-13
 ; PRIOR APPLICATION NUMBER: 60/066770

; PRIOR FILING DATE: 1997-11-24
 ; PRIOR APPLICATION NUMBER: 60/075945
 ; PRIOR FILING DATE: 1998-02-25
 ; PRIOR APPLICATION NUMBER: 60/078910
 ; PRIOR FILING DATE: 1998-03-20
 ; PRIOR APPLICATION NUMBER: 60/083322
 ; PRIOR FILING DATE: 1998-04-28
 ; PRIOR APPLICATION NUMBER: 60/084600
 ; PRIOR FILING DATE: 1998-05-07
 ; PRIOR APPLICATION NUMBER: 60/087106
 ; PRIOR FILING DATE: 1998-05-28
 ; PRIOR APPLICATION NUMBER: 60/087607
 ; PRIOR FILING DATE: 1998-06-02
 ; PRIOR APPLICATION NUMBER: 60/087609
 ; PRIOR FILING DATE: 1998-06-02
 ; PRIOR APPLICATION NUMBER: 60/087759
 ; PRIOR FILING DATE: 1998-06-02
 ; PRIOR APPLICATION NUMBER: 60/087827
 ; PRIOR FILING DATE: 1998-06-03
 ; PRIOR APPLICATION NUMBER: 60/088021
 ; PRIOR FILING DATE: 1998-06-04
 ; PRIOR APPLICATION NUMBER: 60/088025
 ; PRIOR FILING DATE: 1998-06-04
 ; PRIOR APPLICATION NUMBER: 60/088026
 ; PRIOR FILING DATE: 1998-06-04
 ; PRIOR APPLICATION NUMBER: 60/088028
 ; PRIOR FILING DATE: 1998-06-04
 ; PRIOR APPLICATION NUMBER: 60/088029
 ; PRIOR FILING DATE: 1998-06-04
 ; PRIOR APPLICATION NUMBER: 60/088030
 ; PRIOR FILING DATE: 1998-06-04
 ; PRIOR APPLICATION NUMBER: 60/088033
 ; PRIOR FILING DATE: 1998-06-04
 ; PRIOR APPLICATION NUMBER: 60/088326
 ; PRIOR FILING DATE: 1998-06-04
 ; PRIOR APPLICATION NUMBER: 60/088167
 ; PRIOR FILING DATE: 1998-06-05
 ; PRIOR APPLICATION NUMBER: 60/088202
 ; PRIOR FILING DATE: 1998-06-05
 ; PRIOR APPLICATION NUMBER: 60/088212
 ; PRIOR FILING DATE: 1998-06-05
 ; PRIOR APPLICATION NUMBER: 60/088217
 ; PRIOR FILING DATE: 1998-06-05
 ; PRIOR APPLICATION NUMBER: 60/088655
 ; PRIOR FILING DATE: 1998-06-09
 ; PRIOR APPLICATION NUMBER: 60/088734
 ; PRIOR FILING DATE: 1998-06-10
 ; PRIOR APPLICATION NUMBER: 60/088738
 ; PRIOR FILING DATE: 1998-06-10
 ; PRIOR APPLICATION NUMBER: 60/088742
 ; PRIOR FILING DATE: 1998-06-10
 ; PRIOR APPLICATION NUMBER: 60/088810
 ; PRIOR FILING DATE: 1998-06-10
 ; PRIOR APPLICATION NUMBER: 60/088824
 ; PRIOR FILING DATE: 1998-06-10
 ; PRIOR APPLICATION NUMBER: 60/088826
 ; PRIOR FILING DATE: 1998-06-10
 ; PRIOR APPLICATION NUMBER: 60/088858
 ; PRIOR FILING DATE: 1998-06-11
 ; PRIOR APPLICATION NUMBER: 60/088861
 ; PRIOR FILING DATE: 1998-06-11
 ; PRIOR APPLICATION NUMBER: 60/088876
 ; PRIOR FILING DATE: 1998-06-11
 ; PRIOR APPLICATION NUMBER: 60/089105
 ; PRIOR FILING DATE: 1998-06-12
 ; PRIOR APPLICATION NUMBER: 60/089440
 ; PRIOR FILING DATE: 1998-06-16
 ; PRIOR APPLICATION NUMBER: 60/089512
 ; PRIOR FILING DATE: 1998-06-16
 ; PRIOR APPLICATION NUMBER: 60/089514
 ; PRIOR FILING DATE: 1998-06-16
 ; PRIOR APPLICATION NUMBER: 60/089532
 ; PRIOR FILING DATE: 1998-06-17

us-09-270-437d-11.rnpb

Mon Apr 21 10:30:00 2003

;; PRIOR APPLICATION NUMBER: 60/089538
;; PRIOR FILING DATE: 1998-06-17
;; PRIOR APPLICATION NUMBER: 60/089598
;; PRIOR FILING DATE: 1998-06-17
;; PRIOR APPLICATION NUMBER: 60/089599
;; PRIOR FILING DATE: 1998-06-17
;; PRIOR APPLICATION NUMBER: 60/089600
;; PRIOR FILING DATE: 1998-06-17
;; PRIOR APPLICATION NUMBER: 60/089653
;; PRIOR FILING DATE: 1998-06-17
;; PRIOR APPLICATION NUMBER: 60/089801
;; PRIOR FILING DATE: 1998-06-18
;; PRIOR APPLICATION NUMBER: 60/089907
;; PRIOR FILING DATE: 1998-06-18
;; PRIOR APPLICATION NUMBER: 60/089908
;; PRIOR FILING DATE: 1998-06-18
;; PRIOR APPLICATION NUMBER: 60/089947
;; PRIOR FILING DATE: 1998-06-19
;; PRIOR APPLICATION NUMBER: 60/089948
;; PRIOR FILING DATE: 1998-06-19
;; PRIOR APPLICATION NUMBER: 60/089952
;; PRIOR FILING DATE: 1998-06-19
;; PRIOR APPLICATION NUMBER: 60/090246
;; PRIOR FILING DATE: 1998-06-22
;; PRIOR APPLICATION NUMBER: 60/090252
;; PRIOR FILING DATE: 1998-06-22
;; PRIOR APPLICATION NUMBER: 60/090254
;; PRIOR FILING DATE: 1998-06-22
;; PRIOR APPLICATION NUMBER: 60/090349
;; PRIOR FILING DATE: 1998-06-23
;; PRIOR APPLICATION NUMBER: 60/090355
;; PRIOR FILING DATE: 1998-06-23
;; PRIOR APPLICATION NUMBER: 60/090429
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090431
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090435
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090444
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090445
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090472
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090535
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090540
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090542
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090557
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090676
;; PRIOR FILING DATE: 1998-06-25
;; PRIOR APPLICATION NUMBER: 60/090678
;; PRIOR FILING DATE: 1998-06-25
;; PRIOR APPLICATION NUMBER: 60/090690
;; PRIOR FILING DATE: 1998-06-25
;; PRIOR APPLICATION NUMBER: 60/090694
;; PRIOR FILING DATE: 1998-06-25
;; PRIOR APPLICATION NUMBER: 60/090695
;; PRIOR FILING DATE: 1998-06-25
;; PRIOR APPLICATION NUMBER: 60/090696
;; PRIOR FILING DATE: 1998-06-25
;; PRIOR APPLICATION NUMBER: 60/090862
;; PRIOR FILING DATE: 1998-06-26
;; PRIOR APPLICATION NUMBER: 60/090863
;; PRIOR FILING DATE: 1998-06-26
;; PRIOR APPLICATION NUMBER: 60/091360
;; PRIOR FILING DATE: 1998-07-01
;; PRIOR APPLICATION NUMBER: 60/091478
;; PRIOR FILING DATE: 1998-07-02
;; PRIOR APPLICATION NUMBER: 60/091544

;; PRIOR FILING DATE: 1998-07-01
;; PRIOR APPLICATION NUMBER: 60/091519
;; PRIOR FILING DATE: 1998-07-02
;; PRIOR APPLICATION NUMBER: 60/091626
;; PRIOR FILING DATE: 1998-07-02
;; PRIOR APPLICATION NUMBER: 60/091633
;; PRIOR FILING DATE: 1998-07-02
;; PRIOR APPLICATION NUMBER: 60/091978
;; PRIOR FILING DATE: 1998-07-07
;; PRIOR APPLICATION NUMBER: 60/091982
;; PRIOR FILING DATE: 1998-07-07
;; PRIOR APPLICATION NUMBER: 60/092182
;; PRIOR FILING DATE: 1998-07-09

Query Match 74.5%; Score 16.4; DB 9; Length 1759;
Best Local Similarity 94.4%; Pred. No. 73;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 5 TCGCGCTGCTGCGCTCAG 22
Db 98 TCGCGCTGCTGCGCTCAG 115
|||||

RESULT 13
US-09-990-436-180
; Sequence 180, Application US/09990436
; Publication No. US20020198148A1
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi J.
; APPLICANT: Baker, Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Fong, Sherman
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Goddard, Audrey E.
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Kijavini, Ivar J.
; APPLICANT: Napier, Mary A.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE OF INVENTION: Acids Encoding the Same
; FILE REFERENCE: P2730PIC14
; CURRENT APPLICATION NUMBER: US/09/990,436
; CURRENT FILING DATE: 2001-11-14
; PRIOR APPLICATION NUMBER: 60/049787
; PRIOR FILING DATE: 1997-06-16
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/065186
; PRIOR FILING DATE: 1997-11-12
; PRIOR APPLICATION NUMBER: 60/065311
; PRIOR FILING DATE: 1997-11-13
; PRIOR APPLICATION NUMBER: 60/066770
; PRIOR FILING DATE: 1997-11-24
; PRIOR APPLICATION NUMBER: 60/075945
; PRIOR FILING DATE: 1998-02-25
; PRIOR APPLICATION NUMBER: 60/078910
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/083322
; PRIOR FILING DATE: 1998-04-28

/	PRIOR FILING DATE:	1998-06-17
/	PRIOR APPLICATION NUMBER:	60/089653
/	PRIOR FILING DATE:	1998-06-17
/	PRIOR APPLICATION NUMBER:	60/089601
/	PRIOR FILING DATE:	1998-06-18
/	PRIOR APPLICATION NUMBER:	60/089907
/	PRIOR FILING DATE:	1998-06-18
/	PRIOR APPLICATION NUMBER:	60/089908
/	PRIOR FILING DATE:	1998-06-18
/	PRIOR APPLICATION NUMBER:	60/089947
/	PRIOR FILING DATE:	1998-06-19
/	PRIOR APPLICATION NUMBER:	60/089948
/	PRIOR FILING DATE:	1998-06-19
/	PRIOR APPLICATION NUMBER:	60/089952
/	PRIOR FILING DATE:	1998-06-19
/	PRIOR APPLICATION NUMBER:	60/090246
/	PRIOR FILING DATE:	1998-06-22
/	PRIOR APPLICATION NUMBER:	60/090252
/	PRIOR FILING DATE:	1998-06-22
/	PRIOR APPLICATION NUMBER:	60/090254
/	PRIOR FILING DATE:	1998-06-22
/	PRIOR APPLICATION NUMBER:	60/090349
/	PRIOR FILING DATE:	1998-06-23
/	PRIOR APPLICATION NUMBER:	60/090355
/	PRIOR FILING DATE:	1998-06-23
/	PRIOR APPLICATION NUMBER:	60/090429
/	PRIOR FILING DATE:	1998-06-24
/	PRIOR APPLICATION NUMBER:	60/090431
/	PRIOR FILING DATE:	1998-06-24
/	PRIOR APPLICATION NUMBER:	60/090435
/	PRIOR FILING DATE:	1998-06-24
/	PRIOR APPLICATION NUMBER:	60/090444
/	PRIOR FILING DATE:	1998-06-24
/	PRIOR APPLICATION NUMBER:	60/090445
/	PRIOR FILING DATE:	1998-06-24
/	PRIOR APPLICATION NUMBER:	60/090472
/	PRIOR FILING DATE:	1998-06-24
/	PRIOR APPLICATION NUMBER:	60/090535
/	PRIOR FILING DATE:	1998-06-24
/	PRIOR APPLICATION NUMBER:	60/090540
/	PRIOR FILING DATE:	1998-06-24
/	PRIOR APPLICATION NUMBER:	60/090542
/	PRIOR FILING DATE:	1998-06-24
/	PRIOR APPLICATION NUMBER:	60/090557
/	PRIOR FILING DATE:	1998-06-24
/	PRIOR APPLICATION NUMBER:	60/090676
/	PRIOR FILING DATE:	1998-06-25
/	PRIOR APPLICATION NUMBER:	60/090678
/	PRIOR FILING DATE:	1998-06-25
/	PRIOR APPLICATION NUMBER:	60/090690
/	PRIOR FILING DATE:	1998-06-25
/	PRIOR APPLICATION NUMBER:	60/090694
/	PRIOR FILING DATE:	1998-06-25
/	PRIOR APPLICATION NUMBER:	60/090695
/	PRIOR FILING DATE:	1998-06-25
/	PRIOR APPLICATION NUMBER:	60/090696
/	PRIOR FILING DATE:	1998-06-25
/	PRIOR APPLICATION NUMBER:	60/090862
/	PRIOR FILING DATE:	1998-06-26
/	PRIOR APPLICATION NUMBER:	60/090863
/	PRIOR FILING DATE:	1998-06-26
/	PRIOR APPLICATION NUMBER:	60/091360
/	PRIOR FILING DATE:	1998-07-01
/	PRIOR APPLICATION NUMBER:	60/091478
/	PRIOR FILING DATE:	1998-07-02
/	PRIOR APPLICATION NUMBER:	60/091544
/	PRIOR FILING DATE:	1998-07-01
/	PRIOR APPLICATION NUMBER:	60/091519
/	PRIOR FILING DATE:	1998-07-02
/	PRIOR APPLICATION NUMBER:	60/091626
/	PRIOR FILING DATE:	1998-07-02
/	PRIOR APPLICATION NUMBER:	60/091633
/	PRIOR FILING DATE:	1998-07-02

us-09-270-437d-11.rnpb

Mon Apr 21 10:30:00 2003

```

; PRIOR APPLICATION NUMBER: 60/091978
; PRIOR FILING DATE: 1998-07-07
; PRIOR APPLICATION NUMBER: 60/091982
; PRIOR FILING DATE: 1998-07-07
; PRIOR APPLICATION NUMBER: 60/092182
; PRIOR FILING DATE: 1998-07-09

Query Match          74.5%; Score 16.4; DB 9; Length 1759;
Best Local Similarity 94.4%; Pred. No. 73;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 5 TCGCGCTGCGGCTCAG 22
Db 98 TCGCGCTGCTGCTCAG 115

RESULT 14
US-09-991-181-180
; Sequence 180, Application US/09991181
; Publication No. US20020197615A1
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi J.
; APPLICANT: Baker, Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Fong, Sherman
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerlitsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Kijavini, Ivar J.
; APPLICANT: Napier, Mary A.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas P.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; TITLE OF INVENTION: Acids Encoding the Same
; FILE REFERENCE: P27301C53
; CURRENT APPLICATION NUMBER: US/09/991,181
; CURRENT FILING DATE: 2001-11-16
; PRIOR APPLICATION NUMBER: 60/049787
; PRIOR FILING DATE: 1997-06-16
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/065186
; PRIOR FILING DATE: 1997-11-12
; PRIOR APPLICATION NUMBER: 60/065311
; PRIOR FILING DATE: 1997-11-13
; PRIOR APPLICATION NUMBER: 60/066770
; PRIOR FILING DATE: 1997-11-24
; PRIOR APPLICATION NUMBER: 60/075945
; PRIOR FILING DATE: 1998-02-25
; PRIOR APPLICATION NUMBER: 60/078910
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/083322
; PRIOR FILING DATE: 1998-04-28
; PRIOR APPLICATION NUMBER: 60/084600
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/087106
; PRIOR FILING DATE: 1998-05-28
; PRIOR APPLICATION NUMBER: 60/087607
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087609
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/087759
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087827
; PRIOR FILING DATE: 1998-06-03
; PRIOR APPLICATION NUMBER: 60/088021
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088025
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088026
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088028
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088029
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088030
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088033
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088326
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088167
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088202
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088212
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088217
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088655
; PRIOR FILING DATE: 1998-06-09
; PRIOR APPLICATION NUMBER: 60/088734
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088738
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088742
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088810
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088824
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088826
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088858
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/088861
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/088876
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/089105
; PRIOR FILING DATE: 1998-06-12
; PRIOR APPLICATION NUMBER: 60/089440
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089512
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089514
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089532
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089538
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089598
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089599
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089600
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089653
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089801
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/089907
; PRIOR FILING DATE: 1998-06-18

```

; PRIOR APPLICATION NUMBER: 60/089908
 ; PRIOR FILING DATE: 1998-06-18
 ; PRIOR APPLICATION NUMBER: 60/089947
 ; PRIOR FILING DATE: 1998-06-19
 ; PRIOR APPLICATION NUMBER: 60/089948
 ; PRIOR FILING DATE: 1998-06-19
 ; PRIOR APPLICATION NUMBER: 60/089952
 ; PRIOR FILING DATE: 1998-06-19
 ; PRIOR APPLICATION NUMBER: 60/090246
 ; PRIOR FILING DATE: 1998-06-22
 ; PRIOR APPLICATION NUMBER: 60/090252
 ; PRIOR FILING DATE: 1998-06-22
 ; PRIOR APPLICATION NUMBER: 60/090254
 ; PRIOR FILING DATE: 1998-06-22
 ; PRIOR APPLICATION NUMBER: 60/090349
 ; PRIOR FILING DATE: 1998-06-23
 ; PRIOR APPLICATION NUMBER: 60/090355
 ; PRIOR FILING DATE: 1998-06-23
 ; PRIOR APPLICATION NUMBER: 60/090429
 ; PRIOR FILING DATE: 1998-06-24
 ; PRIOR APPLICATION NUMBER: 60/090431
 ; PRIOR FILING DATE: 1998-06-24
 ; PRIOR APPLICATION NUMBER: 60/090435
 ; PRIOR FILING DATE: 1998-06-24
 ; PRIOR APPLICATION NUMBER: 60/090444
 ; PRIOR FILING DATE: 1998-06-24
 ; PRIOR APPLICATION NUMBER: 60/090445
 ; PRIOR FILING DATE: 1998-06-24
 ; PRIOR APPLICATION NUMBER: 60/090472
 ; PRIOR FILING DATE: 1998-06-24
 ; PRIOR APPLICATION NUMBER: 60/090535
 ; PRIOR FILING DATE: 1998-06-24
 ; PRIOR APPLICATION NUMBER: 60/090540
 ; PRIOR FILING DATE: 1998-06-24
 ; PRIOR APPLICATION NUMBER: 60/090542
 ; PRIOR FILING DATE: 1998-06-24
 ; PRIOR APPLICATION NUMBER: 60/090557
 ; PRIOR FILING DATE: 1998-06-24
 ; PRIOR APPLICATION NUMBER: 60/090676
 ; PRIOR FILING DATE: 1998-06-25
 ; PRIOR APPLICATION NUMBER: 60/090678
 ; PRIOR FILING DATE: 1998-06-25
 ; PRIOR APPLICATION NUMBER: 60/090690
 ; PRIOR FILING DATE: 1998-06-25
 ; PRIOR APPLICATION NUMBER: 60/090694
 ; PRIOR FILING DATE: 1998-06-25
 ; PRIOR APPLICATION NUMBER: 60/090695
 ; PRIOR FILING DATE: 1998-06-25
 ; PRIOR APPLICATION NUMBER: 60/090696
 ; PRIOR FILING DATE: 1998-06-25
 ; PRIOR APPLICATION NUMBER: 60/090862
 ; PRIOR FILING DATE: 1998-06-26
 ; PRIOR APPLICATION NUMBER: 60/090863
 ; PRIOR FILING DATE: 1998-06-26
 ; PRIOR APPLICATION NUMBER: 60/091360
 ; PRIOR FILING DATE: 1998-07-01
 ; PRIOR APPLICATION NUMBER: 60/091478
 ; PRIOR FILING DATE: 1998-07-02
 ; PRIOR APPLICATION NUMBER: 60/091544
 ; PRIOR FILING DATE: 1998-07-01
 ; PRIOR APPLICATION NUMBER: 60/091519
 ; PRIOR FILING DATE: 1998-07-02
 ; PRIOR APPLICATION NUMBER: 60/091626
 ; PRIOR FILING DATE: 1998-07-02
 ; PRIOR APPLICATION NUMBER: 60/091633
 ; PRIOR FILING DATE: 1998-07-02
 ; PRIOR APPLICATION NUMBER: 60/091978
 ; PRIOR FILING DATE: 1998-07-07
 ; PRIOR APPLICATION NUMBER: 60/091982
 ; PRIOR FILING DATE: 1998-07-07
 ; PRIOR APPLICATION NUMBER: 60/092182
 ; PRIOR FILING DATE: 1998-07-09

Query Match 74.5%; Score 16.4; DB 9; Length 1759;
 Best Local Similarity 94.4%; Pred. No. 73;
 Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 5 TGGCGCGCTGCGGCTCAG 22
 |||||
 Db 98 TGGCGCGCTGCGGCTCAG 115

RESULT 15

US-09-993-687-180
 ; Sequence 180, Application US/09993687
 ; Publication No. US20020198149A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Ashkenazi, Avi J.
 ; APPLICANT: Baker, Kevin P.
 ; APPLICANT: Botstein, David
 ; APPLICANT: Desnoyers, Luc
 ; APPLICANT: Eaton, Dan L.
 ; APPLICANT: Ferrara, Napoleone
 ; APPLICANT: Fong, Sherman
 ; APPLICANT: Gerber, Hanspeter
 ; APPLICANT: Gerritsen, Mary E.
 ; APPLICANT: Goddard, Audrey
 ; APPLICANT: Godowski, Paul J.
 ; APPLICANT: Grimaldi, J. Christopher
 ; APPLICANT: Gurney, Austin L.
 ; APPLICANT: KJavin, Ivar J.
 ; APPLICANT: Napier, Mary A.
 ; APPLICANT: Pan, James
 ; APPLICANT: Paoni, Nicholas F.
 ; APPLICANT: Roy, Margaret Ann
 ; APPLICANT: Stewart, Timothy A.
 ; APPLICANT: Tumaas, Daniel
 ; APPLICANT: Watanabe, Colin K.
 ; APPLICANT: Williams, P. Mickey
 ; APPLICANT: Wood, William I.
 ; APPLICANT: Zhang, Zemin
 ; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
 ; FILE OF INVENTION: Acids Encoding the Same
 ; FILE REFERENCE: P2730PIC11
 ; CURRENT APPLICATION NUMBER: US/09/993,687
 ; CURRENT FILING DATE: 2002-11-14
 ; PRIOR APPLICATION NUMBER: 60/049787
 ; PRIOR FILING DATE: 1997-06-16
 ; PRIOR APPLICATION NUMBER: 60/062250
 ; PRIOR FILING DATE: 1997-10-17
 ; PRIOR APPLICATION NUMBER: 60/065186
 ; PRIOR FILING DATE: 1997-11-12
 ; PRIOR APPLICATION NUMBER: 60/065311
 ; PRIOR FILING DATE: 1997-11-13
 ; PRIOR APPLICATION NUMBER: 60/066770
 ; PRIOR FILING DATE: 1997-11-24
 ; PRIOR APPLICATION NUMBER: 60/075945
 ; PRIOR FILING DATE: 1998-02-25
 ; PRIOR APPLICATION NUMBER: 60/078910
 ; PRIOR FILING DATE: 1998-03-20
 ; PRIOR APPLICATION NUMBER: 60/083322
 ; PRIOR FILING DATE: 1998-04-28
 ; PRIOR APPLICATION NUMBER: 60/084600
 ; PRIOR FILING DATE: 1998-05-07
 ; PRIOR APPLICATION NUMBER: 60/087106
 ; PRIOR FILING DATE: 1998-05-28
 ; PRIOR APPLICATION NUMBER: 60/087607
 ; PRIOR FILING DATE: 1998-06-02
 ; PRIOR APPLICATION NUMBER: 60/087609
 ; PRIOR FILING DATE: 1998-06-02
 ; PRIOR APPLICATION NUMBER: 60/087759
 ; PRIOR FILING DATE: 1998-06-02
 ; PRIOR APPLICATION NUMBER: 60/087827
 ; PRIOR FILING DATE: 1998-06-03
 ; PRIOR APPLICATION NUMBER: 60/088021
 ; PRIOR FILING DATE: 1998-06-04

PRIOR APPLICATION NUMBER: 60/088025	
PRIOR FILING DATE: 1998-06-04	
PRIOR APPLICATION NUMBER: 60/088026	
PRIOR FILING DATE: 1998-06-04	
PRIOR APPLICATION NUMBER: 60/088028	
PRIOR FILING DATE: 1998-06-04	
PRIOR APPLICATION NUMBER: 60/088029	
PRIOR FILING DATE: 1998-06-04	
PRIOR APPLICATION NUMBER: 60/088030	
PRIOR FILING DATE: 1998-06-04	
PRIOR APPLICATION NUMBER: 60/088033	
PRIOR FILING DATE: 1998-06-04	
PRIOR APPLICATION NUMBER: 60/088326	
PRIOR FILING DATE: 1998-06-04	
PRIOR APPLICATION NUMBER: 60/088167	
PRIOR FILING DATE: 1998-06-05	
PRIOR APPLICATION NUMBER: 60/088202	
PRIOR FILING DATE: 1998-06-05	
PRIOR APPLICATION NUMBER: 60/088212	
PRIOR FILING DATE: 1998-06-05	
PRIOR APPLICATION NUMBER: 60/088217	
PRIOR FILING DATE: 1998-06-05	
PRIOR APPLICATION NUMBER: 60/088655	
PRIOR FILING DATE: 1998-06-09	
PRIOR APPLICATION NUMBER: 60/088734	
PRIOR FILING DATE: 1998-06-10	
PRIOR APPLICATION NUMBER: 60/088738	
PRIOR FILING DATE: 1998-06-10	
PRIOR APPLICATION NUMBER: 60/088742	
PRIOR FILING DATE: 1998-06-10	
PRIOR APPLICATION NUMBER: 60/088810	
PRIOR FILING DATE: 1998-06-10	
PRIOR APPLICATION NUMBER: 60/088824	
PRIOR FILING DATE: 1998-06-10	
PRIOR APPLICATION NUMBER: 60/088826	
PRIOR FILING DATE: 1998-06-10	
PRIOR APPLICATION NUMBER: 60/088858	
PRIOR FILING DATE: 1998-06-11	
PRIOR APPLICATION NUMBER: 60/088861	
PRIOR FILING DATE: 1998-06-11	
PRIOR APPLICATION NUMBER: 60/088876	
PRIOR FILING DATE: 1998-06-11	
PRIOR APPLICATION NUMBER: 60/089105	
PRIOR FILING DATE: 1998-06-12	
PRIOR APPLICATION NUMBER: 60/089440	
PRIOR FILING DATE: 1998-06-16	
PRIOR APPLICATION NUMBER: 60/089512	
PRIOR FILING DATE: 1998-06-16	
PRIOR APPLICATION NUMBER: 60/089514	
PRIOR FILING DATE: 1998-06-16	
PRIOR APPLICATION NUMBER: 60/089532	
PRIOR FILING DATE: 1998-06-17	
PRIOR APPLICATION NUMBER: 60/089538	
PRIOR FILING DATE: 1998-06-17	
PRIOR APPLICATION NUMBER: 60/089598	
PRIOR FILING DATE: 1998-06-17	
PRIOR APPLICATION NUMBER: 60/089599	
PRIOR FILING DATE: 1998-06-17	
PRIOR APPLICATION NUMBER: 60/089600	
PRIOR FILING DATE: 1998-06-17	
PRIOR APPLICATION NUMBER: 60/089653	
PRIOR FILING DATE: 1998-06-17	
PRIOR APPLICATION NUMBER: 60/089801	
PRIOR FILING DATE: 1998-06-18	
PRIOR APPLICATION NUMBER: 60/089907	
PRIOR FILING DATE: 1998-06-18	
PRIOR APPLICATION NUMBER: 60/089908	
PRIOR FILING DATE: 1998-06-19	
PRIOR APPLICATION NUMBER: 60/089947	
PRIOR FILING DATE: 1998-06-19	
PRIOR APPLICATION NUMBER: 60/089948	
PRIOR FILING DATE: 1998-06-19	
PRIOR APPLICATION NUMBER: 60/089952	

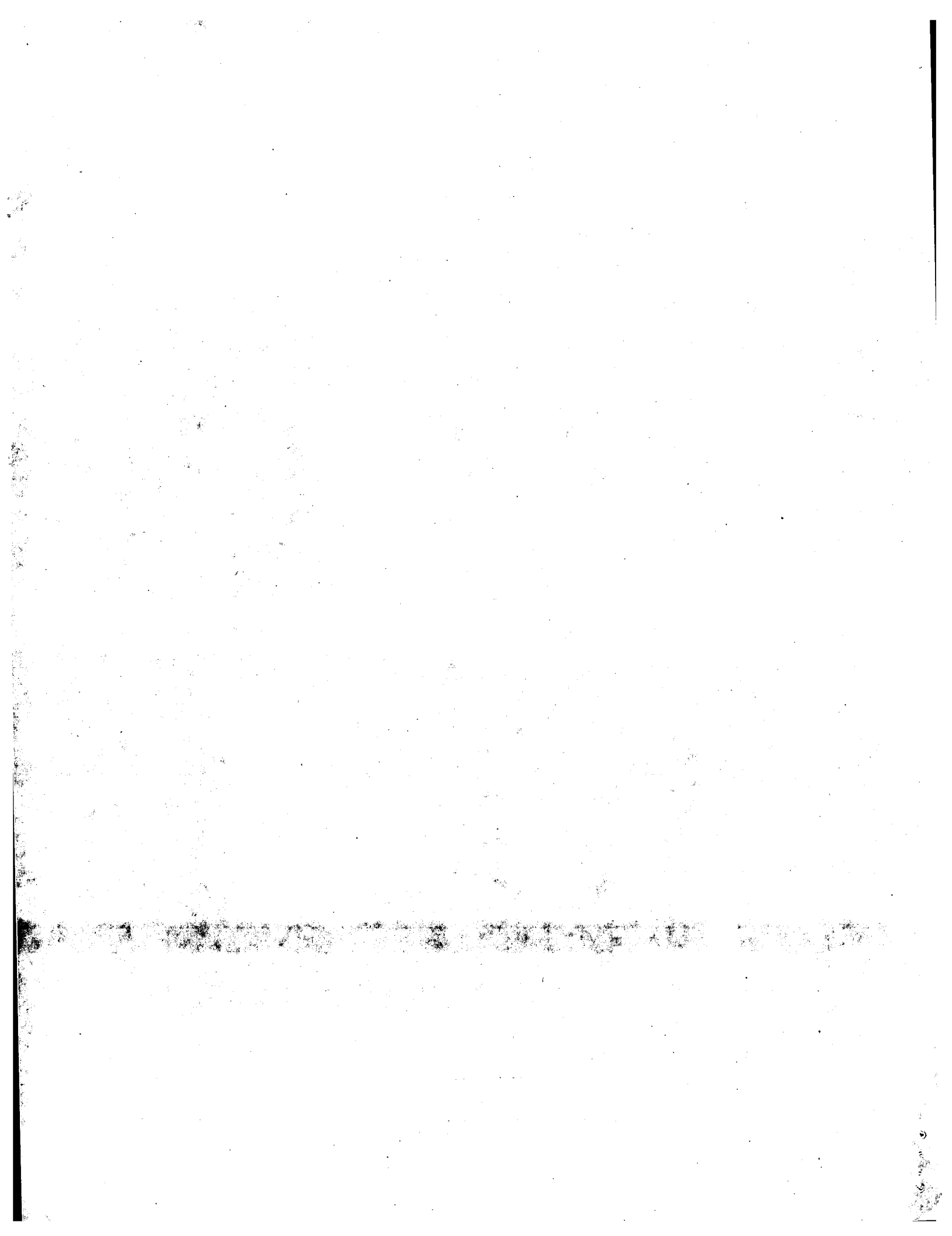
QY	5	TGCGCGCTGCGGCCTCAG	22
pb	98	TGCGCGCTGCTGCCTCAG	115

Mon Apr 21 10:30:00 2003

us-09-270-437d-11.rnpb

Page 15

Search completed: April 18, 2003, 10:14:05
Job time : 94.5 secs



GenCore version 5.1.4.p5 4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 18, 2003, 05:48:17 ; Search time 1211 Seconds
(without alignments)
456.759 Million cell updates/sec

Title: US-09-270-437D-11
Perfect score: 22
Sequence: 1 tcttgcgcgtcgcgcctcag 22

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 24791104 seqs, 12571243825 residues

Total number of hits satisfying chosen parameters: 49582208

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Pending Patents NA Main:

- 1: /cgn2_6/ptodata/1/pna/pctus COMB.seq.*
- 2: /cgn2_6/ptodata/1/pna/US06 COMB.seq.*
- 3: /cgn2_6/ptodata/1/pna/US07 COMB.seq.*
- 4: /cgn2_6/ptodata/1/pna/US08 COMB.seq.*
- 5: /cgn2_6/ptodata/1/pna/US081 COMB.seq.*
- 6: /cgn2_6/ptodata/1/pna/US082 COMB.seq.*
- 7: /cgn2_6/ptodata/1/pna/US083 COMB.seq.*
- 8: /cgn2_6/ptodata/1/pna/US084 COMB.seq.*
- 9: /cgn2_6/ptodata/1/pna/US085 COMB.seq.*
- 10: /cgn2_6/ptodata/1/pna/US086 COMB.seq.*
- 11: /cgn2_6/ptodata/1/pna/US087 COMB.seq.*
- 12: /cgn2_6/ptodata/1/pna/US088 COMB.seq.*
- 13: /cgn2_6/ptodata/1/pna/US089 COMB.seq.*
- 14: /cgn2_6/ptodata/1/pna/US090 COMB.seq.*
- 15: /cgn2_6/ptodata/1/pna/US091 COMB.seq.*
- 16: /cgn2_6/ptodata/1/pna/US092 COMB.seq.*
- 17: /cgn2_6/ptodata/1/pna/US093 COMB.seq.*
- 18: /cgn2_6/ptodata/1/pna/US094 COMB.seq.*
- 19: /cgn2_6/ptodata/1/pna/US095A COMB.seq.*
- 20: /cgn2_6/ptodata/1/pna/US095B COMB.seq.*
- 21: /cgn2_6/ptodata/1/pna/US095C COMB.seq.*
- 22: /cgn2_6/ptodata/1/pna/US095D COMB.seq.*
- 23: /cgn2_6/ptodata/1/pna/US096A COMB.seq.*
- 24: /cgn2_6/ptodata/1/pna/US096B COMB.seq.*
- 25: /cgn2_6/ptodata/1/pna/US096C COMB.seq.*
- 26: /cgn2_6/ptodata/1/pna/US096D COMB.seq.*
- 27: /cgn2_6/ptodata/1/pna/US096E COMB.seq.*
- 28: /cgn2_6/ptodata/1/pna/US096F COMB.seq.*
- 29: /cgn2_6/ptodata/1/pna/US097A COMB.seq.*
- 30: /cgn2_6/ptodata/1/pna/US097B COMB.seq.*
- 31: /cgn2_6/ptodata/1/pna/US097C COMB.seq.*
- 32: /cgn2_6/ptodata/1/pna/US098A COMB.seq.*
- 33: /cgn2_6/ptodata/1/pna/US098B COMB.seq.*
- 34: /cgn2_6/ptodata/1/pna/US098C COMB.seq.*
- 35: /cgn2_6/ptodata/1/pna/US099A COMB.seq.*
- 36: /cgn2_6/ptodata/1/pna/US099B COMB.seq.*
- 37: /cgn2_6/ptodata/1/pna/US099C COMB.seq.*
- 38: /cgn2_6/ptodata/1/pna/US099D COMB.seq.*
- 39: /cgn2_6/ptodata/1/pna/US100A COMB.seq.*
- 40: /cgn2_6/ptodata/1/pna/US100B COMB.seq.*
- 41: /cgn2_6/ptodata/1/pna/US101A COMB.seq.*
- 42: /cgn2_6/ptodata/1/pna/US101B COMB.seq.*
- 43: /cgn2_6/ptodata/1/pna/US102A COMB.seq.*
- 44: /cgn2_6/ptodata/1/pna/US102B COMB.seq.*

- 44: /cgn2_6/ptodata/1/pna/US6000 COMB.seq.*
- 45: /cgn2_6/ptodata/1/pna/US6001 COMB.seq.*
- 46: /cgn2_6/ptodata/1/pna/US6002 COMB.seq.*
- 47: /cgn2_6/ptodata/1/pna/US6003 COMB.seq.*
- 48: /cgn2_6/ptodata/1/pna/US6004 COMB.seq.*
- 49: /cgn2_6/ptodata/1/pna/US6005 COMB.seq.*
- 50: /cgn2_6/ptodata/1/pna/US6006 COMB.seq.*
- 51: /cgn2_6/ptodata/1/pna/US6007 COMB.seq.*
- 52: /cgn2_6/ptodata/1/pna/US6008 COMB.seq.*
- 53: /cgn2_6/ptodata/1/pna/US6009 COMB.seq.*
- 54: /cgn2_6/ptodata/1/pna/US6010 COMB.seq.*
- 55: /cgn2_6/ptodata/1/pna/US6011 COMB.seq.*
- 56: /cgn2_6/ptodata/1/pna/US6012 COMB.seq.*
- 57: /cgn2_6/ptodata/1/pna/US6013 COMB.seq.*
- 58: /cgn2_6/ptodata/1/pna/US6014 COMB.seq.*
- 59: /cgn2_6/ptodata/1/pna/US6015 COMB.seq.*
- 60: /cgn2_6/ptodata/1/pna/US6016 COMB.seq.*
- 61: /cgn2_6/ptodata/1/pna/US6017 COMB.seq.*
- 62: /cgn2_6/ptodata/1/pna/US6018 COMB.seq.*
- 63: /cgn2_6/ptodata/1/pna/US6019 COMB.seq.*
- 64: /cgn2_6/ptodata/1/pna/US6020 COMB.seq.*
- 65: /cgn2_6/ptodata/1/pna/US6021 COMB.seq.*
- 66: /cgn2_6/ptodata/1/pna/US6022 COMB.seq.*
- 67: /cgn2_6/ptodata/1/pna/US6023 COMB.seq.*
- 68: /cgn2_6/ptodata/1/pna/US6024 COMB.seq.*
- 69: /cgn2_6/ptodata/1/pna/US6025 COMB.seq.*
- 70: /cgn2_6/ptodata/1/pna/US6026 COMB.seq.*
- 71: /cgn2_6/ptodata/1/pna/US6027 COMB.seq.*
- 72: /cgn2_6/ptodata/1/pna/US6028 COMB.seq.*
- 73: /cgn2_6/ptodata/1/pna/US6029 COMB.seq.*
- 74: /cgn2_6/ptodata/1/pna/US6030 COMB.seq.*
- 75: /cgn2_6/ptodata/1/pna/US6031 COMB.seq.*
- 76: /cgn2_6/ptodata/1/pna/US6032 COMB.seq.*
- 77: /cgn2_6/ptodata/1/pna/US6033 COMB.seq.*
- 78: /cgn2_6/ptodata/1/pna/US6034 COMB.seq.*
- 79: /cgn2_6/ptodata/1/pna/US6035 COMB.seq.*
- 80: /cgn2_6/ptodata/1/pna/US6036 COMB.seq.*
- 81: /cgn2_6/ptodata/1/pna/US6037 COMB.seq.*
- 82: /cgn2_6/ptodata/1/pna/US6038 COMB.seq.*
- 83: /cgn2_6/ptodata/1/pna/US6039 COMB.seq.*
- 84: /cgn2_6/ptodata/1/pna/US6040 COMB.seq.*
- 85: /cgn2_6/ptodata/1/pna/US6041 COMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	22	100.0	1946	16	US-09-270-437-7	Sequence 7, Appli
2	22	100.0	1946	33	US-09-899-651-7	Sequence 7, Appli
3	18.8	85.5	270	20	US-09-535-896-28782	Sequence 28782, A
C 4	18.8	85.5	525	18	US-09-417-507-8435	Sequence 8435, Ap
C 5	18	81.8	445	27	US-09-696-664A-7808	Sequence 7808, Ap
C 6	18	81.8	445	60	US-60-161-619-7647	Sequence 7647, Ap
C 7	18	81.8	1166	42	US-10-219-999-7854	Sequence 7854, Ap
C 8	17.8	80.9	214	7	US-08-369-881-2595	Sequence 2595, Ap
C 9	17.8	80.9	214	8	US-08-408-872-939	Sequence 939, App
C 10	17.8	80.9	214	20	US-09-535-896-13145	Sequence 13145, A
C 11	17.8	80.9	283	14	US-09-050-817-727	Sequence 727, App
C 12	17.8	80.9	283	16	US-09-270-849B-135409	Sequence 135409, A
C 13	17.8	80.9	283	21	US-09-540-210B-7050	Sequence 7050, Ap
C 14	17.8	80.9	283	48	US-60-043-792-727	Sequence 727, App
C 15	17.8	80.9	301	6	US-08-221-623A-3280	Sequence 3280, Ap
C 16	17.8	80.9	301	6	US-08-221-623B-3280	Sequence 3280, Ap
C 17	17.8	80.9	301	6	US-08-221-623D-3280	Sequence 3280, Ap
C 18	17.8	80.9	432	18	US-09-436-762A-13454	Sequence 13454, A
C 19	17.8	80.9	473	19	US-09-528-409-68226	Sequence 68226, A
C 20	17.8	80.9	473	35	US-09-933-524-68226	Sequence 68226, A
C 21	17.8	80.9	473	35	US-09-933-524A-68226	Sequence 68226, A

ALIGNMENTS

```

? SEQ: 1
? LENGTH: 1946
? TYPE: DNA
? ORGANISM: Homo sapiens
? FEATURE:
US-09-270-437-7

```

Query Match	100.0%;	Score 22;	DB 16;	Length 1946;
Best Local Similarity	100.0%;	Pred. NO. 38;		
Matches 22;	Conservative	0;	Mismatches	0;
			Indels	0;
			Gaps	0;

QY	1	22
QY	TCCTTGGCGGCTGCGGCTCAG	22
db	TCCTTGGCGGCTGCGGCTCAG	59

RESULT 2
US-09-899-651-7
; Sequence 7, Application US/09899651
; GENERAL INFORMATION:
; APPLICANT: Chen, Yao-Tseng
; APPLICANT: Gure, Ali
; APPLICANT: Teang, Solam
; APPLICANT: Stockert, Elisabeth
; APPLICANT: Jager, Elke
; APPLICANT: Knuth, Alexander
; APPLICANT: Old, Lloyd J.

Query Match	100.0%;	Score 22;	DB 33;	Length 1946;
Best Local Similarity	100.0%;	Pred. No. 38;		
Matches	22;	Conservative	0;	Mismatches 0;
				Indels 0

QY 1 TCCTTGGCGGCTGCGGCCTCAG 22
pb 38 TCCTTGGCGGCTGCGGCCTCAG 59

```

RESULT 3
US-09-535-896-28782
; Sequence 28782, Application US/09535896
; GENERAL INFORMATION:
; APPLICANT: Sellhamer, Jeffrey J.
; APPLICANT: Delegeane, Angelo M.
; APPLICANT: Stuart, Susan G.
; APPLICANT: Stuve, Laura L.
; APPLICANT: Mullahy, Sara J.
; APPLICANT: Naughton, Rebecca E.
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING OR REGULATING HYDROLASES
; FILE REFERENCE: PD-1003 CIP
; CURRENT APPLICATION NUMBER: US/09/535,896
; CURRENT FILING DATE: 2000-03-24
; Prior application data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 46268
; SOFTWARE: PERL Program
; SEQ ID NO 28782

```

```

1 ORGANISM: Homo sapiens
2
3 FEATURE:
4
5 NAME/KEY: misc feature
6 OTHER INFORMATION: Incyte ID No: hu00928771
7
8 NAME/KEY: unsure
9
10 LOCATION: 88
11
12 OTHER INFORMATION: a, t, c, g, or other
13
14 TS-09-535-896-28782

```

Query Match	85.5%;	Score 18.8;	DB 20;	Length 270;
Best Local Similarity	90.9%;	Pred. No. 9.1e+02;		
Matches 20;	Conservative	0;	Mismatches 2;	Indels 0

QY
1 TCCTTGGCGGCTGCGGCCTCAG 22

DH
26 TCCTTGGCGGCGCGGGCTCAG 47

RESULT 4
US-09-417-507-8435/c
Sequence 8435, Application US/09417507
GENERAL INFORMATION:
APPLICANT: KEITH G. WEINSTOCK ET AL.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO
ACID SEQUENCES FOR DIAGNOSIS OF
TITLE OF INVENTION: FOMIGRUS FOR DIAGNOSIS
FILE REFERENCE: PATH99-10
CURRENT APPLICATION NUMBER: US/09/417,507
CURRENT FILING DATE: 1999-10-14

```
/ NUMBER OF SEQ ID NOS: 44312
/ SEQ ID NO 8435
/ LENGTH: 525
/ TYPE: DNA
/ ORGANISM: A.fumigatus
US-09-417-507-8435

Query Match      85.5%; Score 18.8; DB 18; Length 525;
Best Local Similarity 90.9%; Pred. No. 8.7e+02;
Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 TCCTTGGCGCTGCGGCCTCAG 22
   |||||
Db 497 TCCTTGGCACTGCGTCTCAG 476

RESULT 5
US-09-696-664A-7808/c
; Sequence 7808, Application US/0969664A
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark S.
; APPLICANT: Andersen, Scott E.
; APPLICANT: Dubois, Patrice
; APPLICANT: Mahadeo, Debbie A.
; APPLICANT: Masucci, James D.
; TITLE OF INVENTION: Nucleic Acid Molecules And Other Molecules Associated With
; FILE REFERENCE: 38-21(51721)B
; CURRENT APPLICATION NUMBER: US/09/696,664A
; CURRENT FILING DATE: 2000-10-25
; PRIOR APPLICATION NUMBER: US 60/161,619
; PRIOR FILING DATE: 1999-10-26
; NUMBER OF SEQ ID NOS: 17472
; SEQ ID NO 7808
; LENGTH: 445
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3279-048-P1-K1-E7
US-09-696-664A-7808

Query Match      81.8%; Score 18; DB 27; Length 445;
Best Local Similarity 100.0%; Pred. No. 1.9e+03;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 CCTTGGCGCTGCGGCCT 19
   |||||
Db 355 CCTTGGCGCTGCGGCCT 338

RESULT 6
US-60-161-619-7647/c
; Sequence 7647, Application US/60161619
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark S.
; APPLICANT: Andersen, Scott E.
; APPLICANT: Dubois, Patrice
; APPLICANT: Mahadeo, Debbie A.
; APPLICANT: Masucci, James D.
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(51721)A
; CURRENT APPLICATION NUMBER: US/60/161,619
; CURRENT FILING DATE: 1999-10-26
; NUMBER OF SEQ ID NOS: 7685
; SEQ ID NO 7647
; LENGTH: 445
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3279-048-P1-K1-E7
US-60-161-619-7647
```

```
Query Match      81.8%; Score 18; DB 60; Length 445;
Best Local Similarity 100.0%; Pred. No. 1.9e+03;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 CCTTGGCGCTGCGGCCT 19
   |||||
Db 355 CCTTGGCGCTGCGGCCT 338

RESULT 7
US-10-219-999-7854/c
; Sequence 7854, Application US/10219999
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Edgerton, Michael D
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Liu, Jingdong
; APPLICANT: Stein, Joshua
; TITLE OF INVENTION: CDNA SEQUENCES AND USES FOR PLANT IMPROVEMENT
; FILE REFERENCE: 38-10(52726)C
; CURRENT APPLICATION NUMBER: US/10/219,999
; CURRENT FILING DATE: 2002-08-15
; PRIOR APPLICATION NUMBER: US 60/324,109
; PRIOR FILING DATE: 2001-09-21
; PRIOR APPLICATION NUMBER: US 60/312,544
; PRIOR FILING DATE: 2001-08-15
; NUMBER OF SEQ ID NOS: 63520
; SEQ ID NO 7854
; LENGTH: 1166
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (2)..(832)
; OTHER INFORMATION:
US-10-219-999-7854

Query Match      81.8%; Score 18; DB 42; Length 1166;
Best Local Similarity 100.0%; Pred. No. 1.8e+03;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 CCTTGGCGCTGCGGCCT 19
   |||||
Db 404 CCTTGGCGCTGCGGCCT 387

RESULT 8
US-08-369-881-2595/c
; Sequence 2595, Application US/08369881
; GENERAL INFORMATION:
; APPLICANT: Wilde, Craig G.
; APPLICANT: Delegeane, Angelo M.
; APPLICANT: Bills, Pamela Kay
; APPLICANT: Pham, Mino Thu;
; TITLE OF INVENTION: HUMAN CARDIAC CELL-DERIVED POLYNUCLEOTIDES
; NUMBER OF SEQUENCES: 3690
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3330 Hillview Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy Disk
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.0/6.1/MS-DOS
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/369,881
; FILING DATE: Herewith
```

US-08-408-872-939

CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: Barbara J. Luther.
REGISTRATION NUMBER: 33,954
REFERENCE/DOCKET NUMBER: PD0019 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-855-0572
INFORMATION FOR SEQ ID NO: 2595:
SEQUENCE CHARACTERISTICS:
LENGTH: 214 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
IMMEDIATE SOURCE:
CLONE: 187978
US-08-369-881-2595

Query Match 80.9%; Score 17.8; DB 7; Length 214;
Best Local Similarity 90.5%; Pred. No. 2.4e+03;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 TCCTTGGCGGCTGGCGGCTCA 21
|||||
Db 157 TCCTTGGCGGCTGGCGGCTCA 137

RESULT 10
US-09-535-896-13145/c
; Sequence 13145, Application US/09535896
; GENERAL INFORMATION:
; APPLICANT: Seilhamer, Jeffrey J.
; APPLICANT: Delegeane, Angelo M.
; APPLICANT: Stuart, Susan G.
; APPLICANT: Stuve, Laura L.
; APPLICANT: Mullaby, Sara J.
; APPLICANT: Naughton, Rebecca E.
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING OR REGULATING HYDROLASES
; FILE REFERENCE: PD-1003 CIP
; CURRENT APPLICATION NUMBER: US/09/535,896
; CURRENT FILING DATE: 2000-03-24
; Prior application data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 46268
; SOFTWARE: PERL Program
; SEQ ID NO 13145
; LENGTH: 214
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No: hu00824962
; NAME/KEY: unsure
; LOCATION: 16, 59, 104, 106, 126, 170, 194, 196, 200
; OTHER INFORMATION: a, t, c, g, or other
US-09-535-896-13145

Query Match 80.9%; Score 17.8; DB 20; Length 214;
Best Local Similarity 90.5%; Pred. No. 2.4e+03;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 TCCTTGGCGGCTGGCGGCTCA 21
|||||
Db 157 TCCTTGGCGGCTGGCGGCTCA 137

RESULT 11
US-09-050-817-727/c
; Sequence 727, Application US/09050817
; GENERAL INFORMATION:
; APPLICANT: Gooding, Douglas H.
; APPLICANT: Stuve, Laura L.
; APPLICANT: Stuart, Susan G.
; APPLICANT: Ito, Laura Y.
; APPLICANT: Akerblom, Ingrid E.
; APPLICANT: Delegeane, Angelo M.
; APPLICANT: Naughton, Rebecca E.
; APPLICANT: Klingler, Tod M.
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM
; TITLE OF INVENTION: HUMAN FETAL KIDNEY
; NUMBER OF SEQUENCES: 2115
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk

```
/
/
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/09/050,817
/ FILING DATE:
/ CLASSIFICATION:
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: 60/043,792
/ FILING DATE:
/ ATTORNEY/AGENT INFORMATION:
/ NAME: CERRONE, MICHAEL C.
/ REGISTRATION NUMBER: 39,132
/ REFERENCE/DOCKET NUMBER: PD-0355P
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (415) 855-0555
/ TELEFAX: (415) 845-4166
/ INFORMATION FOR SEQ ID NO: 727:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 283 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ MOLECULE TYPE: cDNA
/ IMMEDIATE SOURCE:
/ CLONE: 2991071H1
/
/ US-09-050-817-727
/
/ Query Match 80.9%; Score 17.8; DB 14; Length 283;
/ Best Local Similarity 90.5%; Pred. No. 2.4e+03;
/ Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
/
/ QY 2 CCTTGGCGCTGGCGCTCAG 22
/ Db 47 CCTTGGCGCTGCAGCCCG 27
/
/ RESULT 12
/ US-09-270-849B-135409/c
/ Sequence 135409, Application US/09270849B
/ GENERAL INFORMATION:
/ APPLICANT: Swimmer et al.
/ TITLE OF INVENTION: Insect genome survey devices
/ FILE REFERENCE:
/ CURRENT APPLICATION NUMBER: US/09/270,849B
/ CURRENT FILING DATE: 1999-03-17
/ NUMBER OF SEQ ID NOS: 195450
/ SOFTWARE: Patent In Ver. 2.0
/ SEQ ID NO 135409
/ LENGTH: 283
/ TYPE: DNA
/ ORGANISM: Artificial Sequence
/ FEATURE:
/ OTHER INFORMATION: Description of Artificial Sequence: Synthetic
/
/ US-09-270-849B-135409
/
/ Query Match 80.9%; Score 17.8; DB 16; Length 283;
/ Best Local Similarity 90.5%; Pred. No. 2.4e+03;
/ Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
/
/ QY 2 CCTTGGCGCTGGCGCTCAG 22
/ Db 77 CCTTGGCGCTGGCGCTCCG 57
/
/ RESULT 13
/ US-09-540-210B-7050/c
/ Sequence 7050, Application US/09540210B
/ GENERAL INFORMATION:
/ APPLICANT: Sellhammer, Jeffrey J.
/ APPLICANT: Delegeane, Angelo M.
/ APPLICANT: Stuart, Susan G.
/ APPLICANT: Stuve, Laura L.
/
/ APPLICANT: Mullahy, Sara J.
/ APPLICANT: Naughton, Rebecca E.
/ TITLE OF INVENTION: POLYNUCLEOTIDES OF URINARY TRACT TISSUE
/ FILE REFERENCE: PD-1037 CIP
/ CURRENT APPLICATION NUMBER: US/09/540,210B
/ CURRENT FILING DATE: 2002-04-03
/ PRIOR APPLICATION NUMBER: 08/972,899
/ PRIOR FILING DATE: November 18, 1997
/ PRIOR APPLICATION NUMBER: 08/395,244
/ PRIOR FILING DATE: February 27, 1995
/ PRIOR APPLICATION NUMBER: 08/722,922
/ PRIOR FILING DATE: September 27, 1996
/ PRIOR APPLICATION NUMBER: 60/005,526
/ PRIOR FILING DATE: September 23, 1995
/ PRIOR APPLICATION NUMBER: 08/824,029
/ PRIOR FILING DATE: March 25, 1997
/ PRIOR APPLICATION NUMBER: 60/014,010
/ PRIOR FILING DATE: March 25, 1996
/ PRIOR APPLICATION NUMBER: 08/826,847
/ PRIOR FILING DATE: April 10, 1997
/ PRIOR APPLICATION NUMBER: 60/015,533
/ PRIOR FILING DATE: April 10, 1996
/ PRIOR APPLICATION NUMBER: 08/903,555
/ PRIOR FILING DATE: July 31, 1997
/ PRIOR APPLICATION NUMBER: 60/023,308
/ PRIOR FILING DATE: July 31, 1996
/ PRIOR APPLICATION NUMBER: 08/862,178
/ PRIOR FILING DATE: May 22, 1997
/ PRIOR APPLICATION NUMBER: 60/018,217
/ PRIOR FILING DATE: May 23, 1996
/ PRIOR APPLICATION NUMBER: 08/881,589
/ PRIOR FILING DATE: June 24, 1997
/ PRIOR APPLICATION NUMBER: 60/021,275
/ PRIOR FILING DATE: June 25, 1996
/ PRIOR APPLICATION NUMBER: 08/903,802
/ PRIOR FILING DATE: July 31, 1997
/ PRIOR APPLICATION NUMBER: 60/023,308
/ PRIOR FILING DATE: July 31, 1996
/ PRIOR APPLICATION NUMBER: 08/905,881
/ PRIOR FILING DATE: August 1, 1997
/ PRIOR APPLICATION NUMBER: 60/025,204
/ PRIOR FILING DATE: August 1, 1996
/ PRIOR APPLICATION NUMBER: 08/903,471
/ PRIOR FILING DATE: July 30, 1997
/ PRIOR APPLICATION NUMBER: 60/025,478
/ PRIOR FILING DATE: July 31, 1996
/ PRIOR APPLICATION NUMBER: 08/903,556
/ PRIOR FILING DATE: July 31, 1997
/ PRIOR APPLICATION NUMBER: 60/025,217
/ PRIOR FILING DATE: August 22, 1996
/ PRIOR APPLICATION NUMBER: 08/937,142
/ PRIOR FILING DATE: September 23, 1997
/ PRIOR APPLICATION NUMBER: 60/026,598
/ PRIOR FILING DATE: September 24, 1996
/ PRIOR APPLICATION NUMBER: 08/960,746
/ PRIOR FILING DATE: October 29, 1997
/ PRIOR APPLICATION NUMBER: 60/030,144
/ PRIOR FILING DATE: October 30, 1996
/ PRIOR APPLICATION NUMBER: 08/826,847
/ PRIOR FILING DATE: April 10, 1997
/ PRIOR APPLICATION NUMBER: 60/015,533
/ PRIOR FILING DATE: April 10, 1996
/ PRIOR APPLICATION NUMBER: 08/755,524
/ PRIOR FILING DATE: November 22, 1996
/ PRIOR APPLICATION NUMBER: 60/007,495
/ PRIOR FILING DATE: November 22, 1995
/ PRIOR APPLICATION NUMBER: 09/021,031
/ PRIOR FILING DATE: February 10, 1998
/ PRIOR APPLICATION NUMBER: 60/039,325
/ PRIOR FILING DATE: February 13, 1997
/ PRIOR APPLICATION NUMBER: 09/035,172
/ PRIOR FILING DATE: March 4, 1998
/ PRIOR APPLICATION NUMBER: 60/040,431
```

```

; PRIOR FILING DATE: March 5, 1997
; PRIOR APPLICATION NUMBER: 09/041,894
; PRIOR FILING DATE: March 12, 1998
; PRIOR APPLICATION NUMBER: 60/040,199
; PRIOR FILING DATE: March 14, 1997
; PRIOR APPLICATION NUMBER: 09/050,817
; PRIOR FILING DATE: March 30, 1998
; PRIOR APPLICATION NUMBER: 60/043,792
; PRIOR FILING DATE: April 11, 1997
; PRIOR APPLICATION NUMBER: 09/074,999
; PRIOR FILING DATE: May 8, 1998
; PRIOR APPLICATION NUMBER: 60/048,431
; PRIOR FILING DATE: May 29, 1997
; PRIOR APPLICATION NUMBER: 09/107,592
; PRIOR FILING DATE: June 30, 1998
; PRIOR APPLICATION NUMBER: 60/052,751
; PRIOR FILING DATE: July 1, 1997
; PRIOR APPLICATION NUMBER: 09/094,079
; PRIOR FILING DATE: June 9, 1998
; PRIOR APPLICATION NUMBER: 60/049,975
; PRIOR FILING DATE: June 13, 1997
; NUMBER OF SEQ ID NOS: 35634
; SOFTWARE: PERL Program
; SEQ ID NO 7050
; LENGTH: 283
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; OTHER INFORMATION: Incyte ID No: hu00436211
; NAME/KEY: unsure
; LOCATION: 266-267
; OTHER INFORMATION: a, t, c, g, or other
; US-09-540-210B-7050

Query Match 80.9%; Score 17.8; DB 21; Length 283;
Best Local Similarity 90.5%; Pred. No. 2.4e+03;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 CCTTGGCGCTGGCGCTCAG 22
DB 47 CCTTGGCGCTGCAGCCCCAG 27

RESULT 14
US-60-043-792-727/c
; Sequence 727, Application US/60043792
; GENERAL INFORMATION:
; APPLICANT: Gooding, Douglas H.
; APPLICANT: Stuve, Laura L.
; APPLICANT: Stuart, Susan G.
; APPLICANT: Ito, Laura Y.
; APPLICANT: Akerbloom, Ingrid E.
; APPLICANT: Delegeone, Angelo M.
; APPLICANT: Naughton, Rebecca E.
; APPLICANT: Klingler, Tod M.
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM
; TITLE OF INVENTION: HUMAN FETAL KIDNEY
; NUMBER OF SEQUENCES: 2115
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/60/043,792

; PRIOR FILING DATE: March 5, 1997
; PRIOR APPLICATION NUMBER: 09/041,894
; PRIOR FILING DATE: March 12, 1998
; PRIOR APPLICATION NUMBER: 60/040,199
; PRIOR FILING DATE: March 14, 1997
; PRIOR APPLICATION NUMBER: 09/050,817
; PRIOR FILING DATE: March 30, 1998
; PRIOR APPLICATION NUMBER: 60/043,792
; PRIOR FILING DATE: April 11, 1997
; PRIOR APPLICATION NUMBER: 09/074,999
; PRIOR FILING DATE: May 8, 1998
; PRIOR APPLICATION NUMBER: 60/048,431
; PRIOR FILING DATE: May 29, 1997
; PRIOR APPLICATION NUMBER: 09/107,592
; PRIOR FILING DATE: June 30, 1998
; PRIOR APPLICATION NUMBER: 60/052,751
; PRIOR FILING DATE: July 1, 1997
; PRIOR APPLICATION NUMBER: 09/094,079
; PRIOR FILING DATE: June 9, 1998
; PRIOR APPLICATION NUMBER: 60/049,975
; PRIOR FILING DATE: June 13, 1997
; NUMBER OF SEQ ID NOS: 35634
; SOFTWARE: PERL Program
; SEQ ID NO 7050
; LENGTH: 283
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; OTHER INFORMATION: Incyte ID No: hu00436211
; NAME/KEY: unsure
; LOCATION: 266-267
; OTHER INFORMATION: a, t, c, g, or other
; US-09-540-210B-7050

Query Match 80.9%; Score 17.8; DB 21; Length 283;
Best Local Similarity 90.5%; Pred. No. 2.4e+03;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 CCTTGGCGCTGGCGCTCAG 22
DB 47 CCTTGGCGCTGCAGCCCCAG 27

RESULT 15
US-08-221-623A-3280/c
; Sequence 3280, Application US/08221623A
; GENERAL INFORMATION:
; APPLICANT: Adams, Mark D., et. al.
; TITLE OF INVENTION: Human Genes, Sequences, and Expression
; TITLE OF INVENTION: Products
; NUMBER OF SEQUENCES: 3319
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Carella, Byrne, Bain, Gilfillan, Cecchi, Stewart, &
; STREET: 6 Becker Farm Road
; CITY: Roseland
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07068
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 5.0
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/221,623A
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Olstein, Elliot
; REGISTRATION NUMBER: 24,025
; REFERENCE/DOCKET NUMBER: 325800-73
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (201) 994-1700
; TELEFAX: (201) 994-1744
; INFORMATION FOR SEQ ID NO: 3280:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 301 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; US-08-221-623A-3280

Query Match 80.9%; Score 17.8; DB 6; Length 301;

```


Best Local Similarity 90.5%; Pred. No. 2.3e+03;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 TCCTTGGCGGCTGGCGGCTCA 21
|||||
Db 157 TCCTTGGCGGCTGGCGGCTCA 137

Search completed: April 18, 2003, 09:35:44
Job time : 1214 secs

GenCore version 5.1.4 p5 4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 18, 2003, 06:35:28 ; Search time 245.167 Seconds
(without alignments)
400.770 Million cell updates/sec

Title: US-09-270-437D-11

Perfect score: 22

Sequence: 1 tccttgcgcgtcgccctcag 22

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 5897297 seqs, 2233080881 residues

Total number of hits satisfying chosen parameters: 11794594

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Pending Patents NA New.*

- 1: /cgn2_6/ptodata/1/pna/PCT_NEW_COMB.seq.*
- 2: /cgn2_6/ptodata/1/pna/US06_NEW_COMB.seq.*
- 3: /cgn2_6/ptodata/1/pna/US07_NEW_COMB.seq.*
- 4: /cgn2_6/ptodata/1/pna/US08_NEW_COMB.seq.*
- 5: /cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq.*
- 6: /cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq.*
- 7: /cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq.*
- 8: /cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq.*
- 9: /cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq.*
- 10: /cgn2_6/ptodata/1/pna/US60_NEW_COMB.seq.*
- 11: /cgn2_6/ptodata/1/pna/US60_NEW_COMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	22	100.0	22	5	US-09-270-437D-11
2	22	100.0	1946	5	US-09-270-437D-7
3	22	100.0	123600069	5	US-09-948-128-318
4	17.8	80.9	139	6	US-09-513-999C-31413
5	17.8	80.9	301	6	US-09-912-293-219774
6	17.8	80.9	1006	6	PCT-US02-38821-19
7	17.4	79.1	1200	8	US-10-156-761-6708
8	17.4	79.1	15251	7	US-09-949-016-14274
9	17.2	78.2	341	6	US-09-912-293-44484
10	17.2	78.2	423	11	US-60-288-292-2261
11	17.2	78.2	1231	9	US-10-144-771-13059
12	17.2	78.2	1440	9	US-10-366-683-5799
13	17.2	78.2	1920	9	US-10-366-683-5851
14	17.2	78.2	2226	9	US-10-366-683-5851
15	17.2	78.2	1439395	5	US-09-948-128-252
16	17	77.3	609	9	US-10-366-683-10240
17	16.8	76.4	262	6	US-09-912-293-220875
18	16.8	76.4	340	7	US-09-837-604A-1013
19	16.8	76.4	424	7	US-09-837-604A-6577
20	16.8	76.4	468	11	US-60-141-233-62109
21	16.8	76.4	2959	1	PCT-US02-41414-1110
22	16.8	76.4	2976	8	US-10-363-616-152

c	23	16.8	76.4	6866	7	US-09-949-016-15807	Sequence 15807, A
	24	16.4	74.5	214	9	US-10-349-781-38057	Sequence 38057, A
	25	16.4	74.5	1568	9	US-10-144-771-4396	Sequence 4396, Ap
	26	16.4	74.5	1620	6	US-09-962-191A-25	Sequence 25, Appl
	27	16.4	74.5	1620	6	US-09-962-191-25	Sequence 25, Appl
c	28	16.4	74.5	1755	8	US-10-369-493-32171	Sequence 32171, A
	29	16.4	74.5	1759	6	US-09-989-733-180	Sequence 180, App
	30	16.4	74.5	1759	6	US-09-989-733-180	Sequence 180, App
	31	16.4	74.5	1759	9	US-10-125-923A-151	Sequence 151, App
	32	16.4	74.5	1759	9	US-10-205-892-151	Sequence 151, App
	33	16.4	74.5	1759	9	US-10-174-575A-151	Sequence 151, App
	34	16.4	74.5	1759	9	US-10-174-575A-151	Sequence 151, App
	35	16.4	74.5	1759	9	US-10-187-755-151	Sequence 151, App
	36	16.4	74.5	1759	9	US-10-187-749-151	Sequence 151, App
	37	16.4	74.5	1759	9	US-10-199-672-151	Sequence 151, App
	38	16.4	74.5	1759	9	US-10-194-486-151	Sequence 151, App
	39	16.4	74.5	1858	8	US-10-170-235-18030	Sequence 18030, A
	40	16.4	74.5	1889	6	US-09-980-7298-8	Sequence 8, Appl
	41	16.4	74.5	1935	8	US-10-170-235-26844	Sequence 26844, A
	42	16.4	74.5	1974	6	US-09-962-191A-24	Sequence 24, Appl
	43	16.4	74.5	1974	6	US-09-962-191-24	Sequence 24, Appl
c	44	16.2	73.6	208	6	US-09-513-999C-14079	Sequence 14079, A
	45	16.2	73.6	261	7	US-09-675-784A-871	Sequence 871, App

ALIGNMENTS

RESULT 1

US-09-270-437D-11
; Sequence 11, Application US/09270437D

; GENERAL INFORMATION:

; APPLICANT: Chen, Yao-Tseng

; APPLICANT: Gure, Ali

; APPLICANT: Teang, Solam

; APPLICANT: Stockert, Elisabeth

; APPLICANT: Jager, Elke

; APPLICANT: Knuth, Alexander

; APPLICANT: Old, Lloyd J.

; TITLE OF INVENTION: Isolated Nucleic Acid Molecules Encoding Cancer Associated Antigen

; FILE REFERENCE: LUD 5538.1

; CURRENT APPLICATION NUMBER: US/09/270,437D

; CURRENT FILING DATE: 1999-03-16

; PRIOR APPLICATION NUMBER: 09/061,709

; PRIOR FILING DATE: 1998-04-17

; NUMBER OF SEQ ID NOS: 23

; SEQ ID NO 11

; LENGTH: 22

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

US-09-270-437D-11

Query Match 100.0%; Score 22; DB 5; Length 22;

Best Local Similarity 100.0%; Pred. No. 5.2;

Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TCCTTGGCGCTCGCCCTCAG 22

Db 1 TCCTTGGCGCTCGCCCTCAG 22

RESULT 2

US-09-270-437D-7
; Sequence 7, Application US/09270437D

; GENERAL INFORMATION:

; APPLICANT: Chen, Yao-Tseng

; APPLICANT: Gure, Ali

; APPLICANT: Teang, Solam

; APPLICANT: Stockert, Elisabeth

; APPLICANT: Jager, Elke

; APPLICANT: Knuth, Alexander

```

; APPLICANT: Old, Lloyd J.
; TITLE OF INVENTION: Isolated Nucleic Acid Molecules Encoding Cancer Associated Antigens
; FILE REFERENCE: LUD 538.1
; CURRENT APPLICATION NUMBER: US/09/270,437D
; PRIOR FILING DATE: 1999-03-16
; PRIOR APPLICATION NUMBER: 09/061,709
; PRIOR FILING DATE: 1998-04-17
; NUMBER OF SEQ ID NOS: 23
; SEQ ID NO 7
; LENGTH: 1946
; TYPE: DNA
; ORGANISM: Homo sapiens
; NAME/KEY: CDS
; LOCATION: 1622,1702,1771,1809,1833
; OTHER INFORMATION: unsure of nucleotide
US-09-270-437D-7

Query Match      100.0%; Score 22; DB 5; Length 1946;
Best Local Similarity 100.0%; Pred. No. 5.4;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCCTTGGCGCTGGCGCTCAG 22
DB 38 TCCTTGGCGCTGGCGCTCAG 59

RESULT 3
US-09-948-128-318/c
; Sequence 318, Application US/09948128
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig
; TITLE OF INVENTION: COLLECTION OF SINGLE NUCLEOTIDE POLYMORPHISM (SNPs) LOCATED ON CH
; FILE REFERENCE: CL001294
; CURRENT APPLICATION NUMBER: US/09/948,128
; CURRENT FILING DATE: 2001-09-07
; NUMBER OF SEQ ID NOS: 465
; SEQ ID NO 318
; LENGTH: 12360069
; TYPE: DNA
; ORGANISM: HUMAN
; NAME/KEY: misc_feature
; LOCATION: (1)...(12360069)
; OTHER INFORMATION: n = A,T,C or G
US-09-948-128-318

Query Match      100.0%; Score 22; DB 5; Length 12360069;
Best Local Similarity 100.0%; Pred. No. 5.9;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCCTTGGCGCTGGCGCTCAG 22
DB 10941150 TCCTTGGCGCTGGCGCTCAG 10941129

RESULT 4
US-09-513-999C-31413/c
; Sequence 31413, Application US/09513999C
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Duclert, A.
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; FILE REFERENCE: 59.US2.REG
; CURRENT APPLICATION NUMBER: US/09/513,999C
; CURRENT FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/122,487
; PRIOR FILING DATE: 1999-02-26
; NUMBER OF SEQ ID NOS: 36681
; SOFTWARE: Patent.pm

; APPLICANT: Old, Lloyd J.
; TITLE OF INVENTION: Isolated Nucleic Acid Molecules Encoding Cancer Associated Antigens
; FILE REFERENCE: LUD 538.1
; CURRENT APPLICATION NUMBER: US/09/270,437D
; PRIOR FILING DATE: 1999-03-16
; PRIOR APPLICATION NUMBER: 09/061,709
; PRIOR FILING DATE: 1998-04-17
; NUMBER OF SEQ ID NOS: 23
; SEQ ID NO 7
; LENGTH: 1946
; TYPE: DNA
; ORGANISM: Homo sapiens
; NAME/KEY: CDS
; LOCATION: 1622,1702,1771,1809,1833
; OTHER INFORMATION: unsure of nucleotide
US-09-270-437D-7

Query Match      100.0%; Score 22; DB 5; Length 1946;
Best Local Similarity 100.0%; Pred. No. 5.4;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCCTTGGCGCTGGCGCTCAG 22
DB 38 TCCTTGGCGCTGGCGCTCAG 59

RESULT 3
US-09-948-128-318/c
; Sequence 318, Application US/09948128
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig
; TITLE OF INVENTION: COLLECTION OF SINGLE NUCLEOTIDE POLYMORPHISM (SNPs) LOCATED ON CH
; FILE REFERENCE: CL001294
; CURRENT APPLICATION NUMBER: US/09/948,128
; CURRENT FILING DATE: 2001-09-07
; NUMBER OF SEQ ID NOS: 465
; SEQ ID NO 318
; LENGTH: 12360069
; TYPE: DNA
; ORGANISM: HUMAN
; NAME/KEY: misc_feature
; LOCATION: (1)...(12360069)
; OTHER INFORMATION: n = A,T,C or G
US-09-948-128-318

Query Match      100.0%; Score 22; DB 5; Length 12360069;
Best Local Similarity 100.0%; Pred. No. 5.9;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCCTTGGCGCTGGCGCTCAG 22
DB 10941150 TCCTTGGCGCTGGCGCTCAG 10941129

RESULT 4
US-09-513-999C-31413/c
; Sequence 31413, Application US/09513999C
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Duclert, A.
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; FILE REFERENCE: 59.US2.REG
; CURRENT APPLICATION NUMBER: US/09/513,999C
; CURRENT FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/122,487
; PRIOR FILING DATE: 1999-02-26
; NUMBER OF SEQ ID NOS: 36681
; SOFTWARE: Patent.pm
```

```

; SEQ ID NO 31413
; LENGTH: 139
; TYPE: DNA
; ORGANISM: Homo sapiens
; NAME/KEY: misc_feature
; LOCATION: 106
; OTHER INFORMATION: s=g or c
US-09-513-999C-31413

Query Match      80.9%; Score 17.8; DB 6; Length 139;
Best Local Similarity 90.5%; Pred. No. 3.2e+02;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 CCTTGGCGCTGGCGCTCAG 22
DB 50 CCTTGGCGCTGGCGCTCAG 30

RESULT 5
US-09-912-293-219774/c
; Sequence 219774, Application US/09912293
; GENERAL INFORMATION:
; APPLICANT: Rosen, et. al.
; TITLE OF INVENTION: Human Genes, Sequences, and Expression Products 100
; FILE REFERENCE: PO-100
; CURRENT APPLICATION NUMBER: US/09/912,293
; CURRENT FILING DATE: 2001-07-26
; PRIOR APPLICATION NUMBER: 08/103,744
; PRIOR FILING DATE: 1993-08-09
; PRIOR APPLICATION NUMBER: 09/249,651
; PRIOR FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: 08/104,507
; PRIOR FILING DATE: 1993-08-09
; PRIOR APPLICATION NUMBER: 08/196,363
; PRIOR FILING DATE: 1994-02-15
; PRIOR APPLICATION NUMBER: 09/859,490
; PRIOR FILING DATE: 2001-05-18
; PRIOR APPLICATION NUMBER: 08/196,362
; PRIOR FILING DATE: 1994-02-15
; PRIOR APPLICATION NUMBER: 08/221,623
; PRIOR FILING DATE: 1994-03-31
; PRIOR APPLICATION NUMBER: 08/220,691
; PRIOR FILING DATE: 1994-03-31
; PRIOR APPLICATION NUMBER: 09/741,830
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 09/813,155
; PRIOR FILING DATE: 2001-03-21
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 244538
; SEQ ID NO 219774
; LENGTH: 301
; TYPE: DNA
; ORGANISM: Homo sapiens
; NAME/KEY: misc_feature
; LOCATION: (9)..(9)
; OTHER INFORMATION: n is equal to a,t,g, or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (14)..(14)
; OTHER INFORMATION: n is equal to a,t,g, or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (16)..(16)
; OTHER INFORMATION: n is equal to a,t,g, or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (25)..(25)
; OTHER INFORMATION: n is equal to a,t,g, or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (37)..(37)
```

OTHER INFORMATION: n is equal to a,t,g, or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (42)..(42)
OTHER INFORMATION: n is equal to a,t,g, or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (59)..(59)
OTHER INFORMATION: n is equal to a,t,g, or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (61)..(61)
OTHER INFORMATION: n is equal to a,t,g, or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (91)..(91)
OTHER INFORMATION: n is equal to a,t,g, or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (104)..(104)
OTHER INFORMATION: n is equal to a,t,g, or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (126)..(126)
OTHER INFORMATION: n is equal to a,t,g, or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (263)..(263)
OTHER INFORMATION: n is equal to a,t,g, or c
US-09-912-293-219774

Query Match 80.9%; Score 17.8; DB 6; Length 301;
Best Local Similarity 90.5%; Pred. No. 3.2e+02;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 TCCTTGGCGCTCGCGCTCA 21
|||||
DB 157 TCCTTGGCGCTCGCGCTCA 137

RESULT 6
PCT-US02-38821-19/c
Sequence 19, Application PC/TUS0238821
GENERAL INFORMATION:
APPLICANT: CuroGen Corp. et al.
TITLE OF INVENTION: NOVEL HUMAN PROTEINS, POLYNUCLEOTIDES ENCODING THEM AND METHODS
FILE REFERENCE: 24102-502D-061
CURRENT APPLICATION NUMBER: PCT/US02/38821
PRIOR FILING DATE: 2002-12-03
PRIOR APPLICATION NUMBER: 60/341,477
PRIOR FILING DATE: 2001-12-17
PRIOR APPLICATION NUMBER: 60/381,495
PRIOR FILING DATE: 2002-05-17
PRIOR APPLICATION NUMBER: 60/401,788
PRIOR FILING DATE: 2002-08-07
PRIOR APPLICATION NUMBER: 60/341,540
PRIOR FILING DATE: 2001-12-17
PRIOR APPLICATION NUMBER: 60/383,744
PRIOR FILING DATE: 2002-05-28
PRIOR APPLICATION NUMBER: 60/342,592
PRIOR FILING DATE: 2001-12-20
PRIOR APPLICATION NUMBER: 60/344,903
PRIOR FILING DATE: 2001-12-31
PRIOR APPLICATION NUMBER: 60/384,204
PRIOR FILING DATE: 2002-05-29
PRIOR APPLICATION NUMBER: 60/373,288
PRIOR FILING DATE: 2002-04-17
PRIOR APPLICATION NUMBER: 60/380,981
PRIOR FILING DATE: 2002-05-15
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 53
SOFTWARE: CuroSeqList version 0.1

SEQ ID NO 19
LENGTH: 1006
TYPE: DNA
ORGANISM: CG164521-01
FEATURE:
NAME/KEY: CDS
LOCATION: (24)..(995)
PCT-US02-38821-19

Query Match 80.9%; Score 17.8; DB 1; Length 1006;
Best Local Similarity 90.5%; Pred. No. 3.2e+02;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 2 CCTTGGCGCTCGCGCTCAG 22
|||||
DB 50 CCTTGGCGCTCGCGCTCAG 30

RESULT 7
US-10-156-761-6708/c
Sequence 6708, Application US/10156761
GENERAL INFORMATION:
APPLICANT: OMURA, SATOSHI
APPLICANT: IKEDA, HARUO
APPLICANT: ISHIKAWA, JUN
APPLICANT: HORIKAWA, HIROSHI
APPLICANT: SHIBA, TADAYOSHI
APPLICANT: SAKAKI, YOSHIYUKI
APPLICANT: HATTORI, MASAHIRA
TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
FILE REFERENCE: 249-262
CURRENT APPLICATION NUMBER: US/10/156,761
CURRENT FILING DATE: 2002-05-29
PRIOR APPLICATION NUMBER: JP 2001-204089
PRIOR FILING DATE: 2001-05-30
PRIOR APPLICATION NUMBER: JP 2001-272697
PRIOR FILING DATE: 2001-08-02
NUMBER OF SEQ ID NOS: 15109
SEQ ID NO 6708
LENGTH: 1200
TYPE: DNA
ORGANISM: Streptomyces avermitilis
FEATURE:
NAME/KEY: CDS
LOCATION: (1)..(1200)
US-10-156-761-6708

Query Match 79.1%; Score 17.4; DB 8; Length 1200;
Best Local Similarity 94.7%; Pred. No. 4.8e+02;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 TCCTTGGCGCTCGCGCT 19
|||||
DB 668 TCCTTGGCGCTCGCGCT 650

RESULT 8
US-09-949-016-14274
Sequence 14274, Application US/09949016
GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: CL001307
CURRENT APPLICATION NUMBER: US/09/949,016
CURRENT FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012

us-09-270-437d-11.rnpn

Mon Apr 21 10:30:03 2003

```

; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14274
; LENGTH: 15251
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-14274

Query Match      79.1%; Score 17.4; DB 7; Length 15251;
Best Local Similarity 94.7%; Pred. No. 4.9e+02;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 CCTTGGCGGCTGGCGGCTC 20
DB 116 CCTTGGCGGCTGGCGGCTC 134

RESULT 9
US-09-912-293-44484/c
; Sequence 44484, Application US/0912293
; GENERAL INFORMATION:
; APPLICANT: Rosen, et. al.
; TITLE OF INVENTION: Human Genes, Sequences, and Expression Products 100
; FILE REFERENCE: PO-100
; CURRENT APPLICATION NUMBER: US/09/912,293
; CURRENT FILING DATE: 2001-07-26
; PRIOR APPLICATION NUMBER: 08/103,744
; PRIOR FILING DATE: 1993-08-09
; PRIOR APPLICATION NUMBER: 09/249,651
; PRIOR FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: 08/104,507
; PRIOR FILING DATE: 1993-08-09
; PRIOR APPLICATION NUMBER: 08/196,363
; PRIOR FILING DATE: 1994-02-15
; PRIOR APPLICATION NUMBER: 09/859,490
; PRIOR FILING DATE: 2001-05-18
; PRIOR APPLICATION NUMBER: 08/196,362
; PRIOR FILING DATE: 1994-02-15
; PRIOR APPLICATION NUMBER: 08/221,623
; PRIOR FILING DATE: 1994-03-31
; PRIOR APPLICATION NUMBER: 08/220,691
; PRIOR FILING DATE: 1994-03-31
; PRIOR APPLICATION NUMBER: 09/741,830
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 09/813,155
; PRIOR FILING DATE: 2001-03-21
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 244538
; SEQ ID NO 44484
; LENGTH: 341
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (236)..(236)
; OTHER INFORMATION: n is equal to a,t,g, or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (257)..(257)
; OTHER INFORMATION: n is equal to a,t,g, or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (280)..(280)
; OTHER INFORMATION: n is equal to a,t,g, or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (285)..(285)
; OTHER INFORMATION: n is equal to a,t,g, or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (326)..(326)
; OTHER INFORMATION: n is equal to a,t,g, or c
; FEATURE:
; NAME/KEY: misc_feature

```

```

; LOCATION: (333)..(333)
; OTHER INFORMATION: n is equal to a,t,g, or c
US-09-912-293-44484

Query Match      78.2%; Score 17.2; DB 6; Length 341;
Best Local Similarity 86.4%; Pred. No. 5.8e+02;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 TCCTTGGCGGCTGGCGGCTCAG 22
DB 200 TCCTTGGCGGCTGGCGGCTCAG 179

RESULT 10
US-60-288-292-2261/c
; Sequence 2261, Application US/60288292
; GENERAL INFORMATION:
; APPLICANT: Glenn, Matthew
; APPLICANT: Norris, Michael G
; TITLE OF INVENTION: Compounds isolated from forage plants
; TITLE OF INVENTION: and methods for their use.
; FILE REFERENCE: 1058P
; CURRENT APPLICATION NUMBER: US/60/288,292
; CURRENT FILING DATE: 2001-05-02
; NUMBER OF SEQ ID NOS: 49762
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2261
; LENGTH: 423
; TYPE: DNA
; ORGANISM: Lolium perenne
US-60-288-292-2261

Query Match      78.2%; Score 17.2; DB 11; Length 423;
Best Local Similarity 86.4%; Pred. No. 5.8e+02;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 TCCTTGGCGGCTGGCGGCTCAG 22
DB 288 TCCTTGGCGGCTGGCGGCTCAG 267

RESULT 11
US-10-144-771-13059
; Sequence 13059, Application US/10144771
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig
; TITLE OF INVENTION: HUMAN GENOME DISCOVERY SYSTEM AND USES THEREOF
; FILE REFERENCE: CLO01321
; CURRENT APPLICATION NUMBER: US/10/144,771
; CURRENT FILING DATE: 2002-05-15
; NUMBER OF SEQ ID NOS: 47235
; SEQ ID NO 13059
; LENGTH: 1231
; TYPE: DNA
; ORGANISM: HUMAN
US-10-144-771-13059

Query Match      78.2%; Score 17.2; DB 9; Length 1231;
Best Local Similarity 86.4%; Pred. No. 5.8e+02;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 TCCTTGGCGGCTGGCGGCTCAG 22
DB 357 TCCTTGGCGGCTGGCGGCTCAG 378

RESULT 12
US-10-366-683-5799
; Sequence 5799, Application US/10366683
; GENERAL INFORMATION:
; APPLICANT: Rubenfield, Marc J.
; APPLICANT: Nolling, Jork
; APPLICANT: Deloughery, Craig

```

```
; APPLICANT: Bush, David
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; CURRENT APPLICATION NUMBER: US/10/366,683
; CURRENT FILING DATE: 2003-02-13
; PRIOR APPLICATION NUMBER: 09/252,991
; PRIOR FILING DATE: 1999-02-18
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 5799
; LENGTH: 1440
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-10-366-683-5799

Query Match
Best Local Similarity 78.2%; Score 17.2; DB 9; Length 1440;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 TCCTTGGCGGCTCGGCCTCAG 22
|||||
DB 1301 TCCTTGGCGGCGGCTCCTCGG 1322

RESULT 13
US-10-366-683-5851/c
; Sequence 5851, Application US/10366683
; GENERAL INFORMATION:
; APPLICANT: Rubenfield, Marc J.
; APPLICANT: Nollong, Jork
; APPLICANT: Deloughery, Craig
; APPLICANT: Bush, David
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: PATH03-04
; CURRENT APPLICATION NUMBER: US/10/366,683
; CURRENT FILING DATE: 2003-02-13
; PRIOR APPLICATION NUMBER: 09/252,991
; PRIOR FILING DATE: 1999-02-18
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 5851
; LENGTH: 1920
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-10-366-683-5851

Query Match
Best Local Similarity 78.2%; Score 17.2; DB 9; Length 1920;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 TCCTTGGCGGCTCGGCCTCAG 22
|||||
DB 174 TCCTTGGCGGCGGCTCCTCGG 153

RESULT 14
US-10-366-683-5869/c
; Sequence 5869, Application US/10366683
; GENERAL INFORMATION:
; APPLICANT: Rubenfield, Marc J.
; APPLICANT: Nollong, Jork
; APPLICANT: Deloughery, Craig
; APPLICANT: Bush, David
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: PATH03-04
; CURRENT APPLICATION NUMBER: US/10/366,683
; CURRENT FILING DATE: 2003-02-13
; PRIOR APPLICATION NUMBER: 09/252,991
; PRIOR FILING DATE: 1999-02-18
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 5869
; LENGTH: 2226
```

```
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-10-366-683-5869
```

```
Query Match
Best Local Similarity 78.2%; Score 17.2; DB 9; Length 2226;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
```

```
QY 1 TCCTTGGCGGCTCGGCCTCAG 22
|||||
DB 143 TCCTTGGCGGCGGCTCCTCGG 122
```

RESULT 15

```
US-09-948-128-252/c
; Sequence 252, Application US/09948128
```

```
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig
; TITLE OF INVENTION: COLLECTION OF SINGLE NUCLEOTIDE POLYMORPHISM (SNPs) LOCATED ON CHF
; FILE REFERENCE: CL001294
; CURRENT APPLICATION NUMBER: US/09/948,128
```

```
; CURRENT FILING DATE: 2001-09-07
; NUMBER OF SEQ ID NOS: 465
; SEQ ID NO 252
```

```
; LENGTH: 1439395
; TYPE: DNA
```

```
; ORGANISM: HUMAN
```

```
; FEATURE:
```

```
; NAME/KEY: misc feature
```

```
; LOCATION: (1)...(1439395)
```

```
; OTHER INFORMATION: n = A,T,C or G
```

```
US-09-948-128-252
```

```
Query Match
Best Local Similarity 78.2%; Score 17.2; DB 5; Length 1439395;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
```

```
QY 1 TCCTTGGCGGCTCGGCCTCAG 22
|||||
```

```
DB 1321142 TCATGAGCTTCGCGCCTCAG 1321121
```

```
Search completed: April 18, 2003, 10:50:26
Job time : 3044.17 secs
```


GenCore version 5.1.4 p5 4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic-- nucleic search, using sw model

Run on: April 18, 2003, 04:47:40 ; Search time 766 Seconds
(without alignments)
465.145 Million cell updates/sec

Title: US-09-270-437D-11

Perfect score: 22
Sequence: 1 tccttgccgctgcggcctcag 22

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST.*

1: em_estba.*

2: em_estum.*

3: em_estin.*

4: em_estmu.*

5: em_estov.*

6: em_estpl.*

7: em_estro.*

8: em_hcc.*

9: gb_estl.*

10: gb_est2.*

11: gb_hcc.*

12: gb_est3.*

13: gb_est4.*

14: gb_est5.*

15: em_estfun.*

16: em_estom.*

17: gb_gss.*

18: em_gss_hum.*

19: em_gss_inv.*

20: em_gss_pln.*

21: em_gss_vrt.*

22: em_gss_fun.*

23: em_gss_mam.*

24: em_gss_mue.*

25: em_gss_other.*

26: em_gss_pro.*

27: em_gss_rod.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	18.4	83.6	1055	17	CNS01EYE
C 2	18.4	83.6	1128	17	BQ954241
C 3	18.4	83.6	1696	12	BF312525
C 4	17.8	80.9	250	9	AA591524
C 5	17.8	80.9	448	9	AA433531
C 6	17.8	80.9	461	10	AW052407

C 7	17.8	80.9	490	12	BG465076
C 8	17.8	80.9	512	10	AV433340
C 9	17.8	80.9	708	10	BE254482
C 10	17.8	80.9	767	12	BG599633
C 11	17.8	80.9	856	9	AL519496
C 12	17.8	80.9	905	9	AL528423
C 13	17.8	80.9	936	17	CNS035RV
C 14	17.8	80.9	938	14	BQ684757
C 15	17.8	80.9	961	12	BG617183
C 16	17.8	80.9	1013	12	BG757407
C 17	17.4	79.1	432	9	AA120169
C 18	17.4	79.1	862	13	BI818151
C 19	17.2	78.2	291	9	AL825604
C 20	17.2	78.2	319	12	BF886876
C 21	17.2	78.2	367	10	BF843150
C 22	17.2	78.2	382	10	BB842521
C 23	17.2	78.2	437	10	BB864106
C 24	17.2	78.2	460	9	AA000295
C 25	17.2	78.2	497	17	BH035572
C 26	17.2	78.2	508	10	BB858287
C 27	17.2	78.2	512	10	BE544739
C 28	17.2	78.2	546	10	AW210189
C 29	17.2	78.2	552	12	BG086730
C 30	17.2	78.2	552	14	BQ606886
C 31	17.2	78.2	594	17	AZ879071
C 32	17.2	78.2	597	12	BF719752
C 33	17.2	78.2	600	12	BG086847
C 34	17.2	78.2	603	10	BB630902
C 35	17.2	78.2	613	17	BH886739
C 36	17.2	78.2	620	9	AA797109
C 37	17.2	78.2	623	13	BI077950
C 38	17.2	78.2	644	10	BB613343
C 39	17.2	78.2	644	10	BE291303
C 40	17.2	78.2	672	13	BI659501
C 41	17.2	78.2	705	10	BE414309
C 42	17.2	78.2	710	12	BF786998
C 43	17.2	78.2	713	17	BH577830
C 44	17.2	78.2	776	17	AQ745151
C 45	17.2	78.2	800	17	CNS06YME

ALIGNMENTS

RESULT 1	CNS01EYE/c	1055 bp	DNA	linear	GSS 30-MAY-2001
LOCUS	Anopheles gambiae GSS SP6 end of clone 03E18 of NotreDame library				
DEFINITION	from strain PEST of Anopheles gambiae (African malaria mosquito), genomic survey sequence.				
ACCESSION	AL141159				
VERSION	AL141159.1				GI:6999277
KEYWORDS	GSS:				
SOURCE	African malaria mosquito.				
ORGANISM	Anopheles gambiae				
	Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae; Anopheles.				
REFERENCE	1 (bases 1 to 1055)				
AUTHORS	Genoscope.				
TITLE	Direct Submission				
JOURNAL	Submitted (16-FEB-2000) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)				
REFERENCE	2 (bases 1 to 1055)				
AUTHORS	Roth, C.W., Brey, P.T., Ke, Z., Collins, F.H. and Weissenbach, J.				
TITLE	Direct Submission				
JOURNAL	Submitted (16-FEB-2000) BBMI, Institut Pasteur, 25, rue du Dr. Roux, Paris 75015, France				
COMMENT	This clone is from an A. gambiae BAC library provided by F.H. Collins and sequenced by Genoscope in collaboration with the Laboratory of Biochem. and Biol. Molec. of Insects, Institut Pasteur.				

```

FEATURES
source
Location/Qualifiers
1. .1055
/organism="Anopheles gambiae"
/strain="PEST"
/db_xref="taxon:7165"
/clone="03E18"
/clone_lib="NotreDamel"
/notes="end : SP6"
BASE COUNT 244 a 230 c 255 g 320 t 6 others
ORIGIN
Query Match 83.6%; Score 18.4; DB 17; Length 1055;
Best Local Similarity 95.0%; Pred. No. 1.2e+03;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 CCTTGGCGCGCTGGCGCTCA 21
|||||
Db 262 CCTTGGCGCGCTGGCGCTCA 243

RESULT 2
BQ954241 1128 bp mRNA linear EST 21-AUG-2002
AGENCOURT 8864349 lupski_sciatic_nerve Homo sapiens cDNA clone
IMAGE:6197933 5', mRNA sequence.
ACCESSION BQ954241
VERSION BQ954241.1 GI:22369719
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1128)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgaabs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM13608 row: m column: 06
High quality sequence stop: 338.
FEATURES
source
Location/Qualifiers
1. .1128
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:6197933"
/clone_lib="Lupski_sciatic_nerve"
/sex="male"
/tissue type="sciatic nerve"
/dev stage="adult, 70 yr"
/lab host="DH10B"
/notes="Vector: PCMV-SPORT6 (Life Technologies); Site_1:
NotI; Site_2: SalI; cDNA made by oligo-dt priming.
Directionally cloned using the following adaptors:
5'-GACTAGTTCATGACGAGCGCGCCCT(15)-3'. Size selected >
1 kb for average insert length 1.87 kb. This is a primary
library, non-amplified. Library constructed by Life
Technologies and donated by J. Lupski, M.D./Ph.D. (Baylor
College of Medicine) and is available through Life
Technologies."
BASE COUNT 296 a 324 c 322 g 186 t
ORIGIN
Query Match 83.6%; Score 18.4; DB 14; Length 1128;
Best Local Similarity 95.0%; Pred. No. 1.2e+03;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

FEATURES
source
Location/Qualifiers
1. .1696
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4126750"
/clone_lib="NIH MGC 19"
/tissue type="neuroblastoma"
/lab host="DH10B (phage-resistant)"
/notes="Organ: brain; Vector: pOTS7; Site_1: XhoI; Site_2:
EcoRI; cDNA made by oligo-dt priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adaptor: GGCACGAG(G). Library constructed by Ling Hong
in the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH MGC Library."
BASE COUNT 420 a 395 c 535 g 346 t
ORIGIN
Query Match 83.6%; Score 18.4; DB 12; Length 1696;
Best Local Similarity 95.0%; Pred. No. 1.3e+03;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TCCTTGGCGCGCTGGCGCTC 20
|||||
Db 1315 TCCTTGGCGCGCTGGCGCTC 1296

RESULT 4
AA591524 250 bp mRNA linear EST 16-SEP-1997
LOCUS vk23e01.r1 Knowles Solter mouse blastocyst B1 Mus musculus CDNA
DEFINITION clone IMAGE:947448 5', mRNA sequence.
ACCESSION AA591524
VERSION AA591524.1 GI:2404837
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 250)

```

AUTHORS

Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T., Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M., Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B., Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and Waterston,R.

TITLE
JOURNAL
COMMENT

The WashU-HHMI Mouse EST Project
Unpublished (1996)
Contact: Marra M/Mouse EST Project
WashU-HHMI Mouse EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810

Email: mouseest@wustl.edu
This clone is available royalty-free through LLNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:544304

Putative full length read
vector to vector length is 251.

FEATURES
source

Location/Qualifiers
1. .250

/organism="Mus musculus"
/strain="C57BL/6J x DBA/2J F1"
/db_xref="taxon:10090"
/clone="IMAGE:947448"
/clone_lib="Knowles Solter mouse blastocyst B1"
/tissue_type="blastocyst"
/dev_stage="embryo (pre-implantation)"
/lab_host="DH10B"

/note="Organ: embryo; Vector: pSPORT; Site1: NotI; Site 2: SalI; Cloned unidirectionally from mRNA prepared from 800 blastocysts. Primer: SalI(dT): 5'-CGTGCACGCTGCGCTTTT-3'. cDNAs were cloned into the NotI/SalI sites of a pSPORT vector (Life Technologies). Two different size selections: B1 (larger inserts) and B3."

BASE COUNT
ORIGIN

55 a 80 c 82 g 33 t

Query Match 80.9%; Score 17.8; DB 9; Length 250;
Best Local Similarity 90.5%; Pred. NO. 1.6e+03;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 CCTTGGCGCTCGGCTCAG 22
|||
Db 151 CCTTGGCGCTCGGCTCAG 131
|||

RESULT 5
AA433531/c

LOCUS
DEFINITION
vfa3a06.r1 Soares mouse NBMH Mus musculus cDNA clone IMAGE:846514
5', mRNA sequence.

ACCESSION
VERSION
AA433531.1 GI:2138449

KEYWORDS
SOURCE

house mouse.

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 448)
Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T., Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M., Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B., Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and Waterston,R.

The WashU-HHMI Mouse EST Project
Unpublished (1996)

Contact: Marra M/Mouse EST Project
WashU-HHMI Mouse EST Project

Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108

Tel: 314 286 1800

Fax: 314 286 1810

Email: mouseest@wustl.edu

This clone is available royalty-free through LLNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:498666

Seq primer: -28ml3 rev2 ET from Amersham

High quality sequence stop: 439.

FEATURES
source

Location/Qualifiers
1. .448

/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="IMAGE:846514"
/clone_lib="Soares mouse NBMH"
/sex="male"
/tissue_type="heart"
/dev_stage="4 weeks"
/lab_host="DH10B"

/note="Vector: pT73D-Pac (Pharmacia) with a modified polylinker; Site1: Not I; Site 2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5'-TGTTACCAATCTGAAGTGGAGCGCGCGAAGTTT-3']; double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT73 vector. RNA provided by Dr. Minoru Ko, Wayne State Univ. Library constructed and normalized by Bento Soares and M.Patima Bonaudo."

BASE COUNT 90 a 143 c 115 g 100 t

ORIGIN

Query Match 80.9%; Score 17.8; DB 9; Length 448;
Best Local Similarity 90.5%; Pred. NO. 1.8e+03;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 CCTTGGCGCTCGGCTCAG 22
|||

Db 210 CCATGGCGCTCGGCTCAG 190
|||

RESULT 6
AW052407/c

LOCUS

rs35b02.y1 Sommer Pristionchus Pristionchus pacificus cDNA 5', similar to SM:MYSA_CAEEL P12844 MYOSIN HEAVY CHAIN A 1, mRNA sequence.

ACCESSION
VERSION
AW052407.1 GI:5914766

KEYWORDS
SOURCE

Pristionchus pacificus.

ORGANISM

Eukaryota; Metazoa; Nematoda; Chromadorea; Diplogasterida;

REFERENCE
AUTHORS

1 (bases 1 to 461)
McCarte,J., Clifton,S., Chiapelli,B., Pape,D., Martin,J., Wylie,T., Dante,M., Marra,M., Hillier,L., Kucaba,T., Theising,B., Bowers,Y., Gibbons,M., Ritter,E., Bennett,J., Franklin,C., Tsagareishvili,R., Ronko,I., Kennedy,S., Maguire,L., Beck,C., Underwood,K., Steptoe,M., Allen,M., Person,B., Swaller,T., Harvey,N., Schurk,R., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R., Waterston,R. and Wilson,R.

The Washington Univ. Nematode EST Project, 1999
Unpublished (1999)

Contact: McCarter JP

The Washington Univ. Nematode EST Project, 1999
Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810

Email: est@wustl.edu

The library was constructed by Dr. Ralf Sommer DNA Sequencing by:
Washington University Genome Sequencing Center
Contact Dr. Ralf Sommer (ralf.sommer@wustl.edu) for

us-09-270-437d-11.rst

Mon Apr 21 10:30:04 2003

information about this clone.
Seq primer: -40RP from Gibco
High quality sequence stop: 414.

FEATURES

source
1. .461
/organism="Pristionchus pacificus"
/db_xref="taxon:54126"
/clone_lib="Sommer Pristionchus"
/sex="predominantly hermaphroditic"
/dev_stage="mixed stages (embryo to adult)"
/lab_host="not applicable (host cell line)"
/note="vector: Uni-ZAP XR Vector (Stratagene); Site_1: 5'
EcoRI; Site_2: 3' XhoI; 1st strand cDNA was primed with a
XhoI - oligo(dT) primer. Double-stranded cDNA was ligated
to EcoRI adaptors digested with XhoI and cloned into XhoI
and EcoRI sites. Primary complexity of the library was 10
in the 7th. The library went through one round of
amplification."

BASE COUNT 135 a 118 c 154 g 54 t
ORIGIN

Query Match 80.9%; Score 17.8; DB 10; Length 461;
Best Local Similarity 90.5%; Pred. No. 1.8e+03;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 CCTTGGCGCTGGCGCTCAG 22

|||||
Db 217 CCTTGGCGCTGGCGCTCAG 197

RESULT 7

LOCUS BG465076/c 490 bp mRNA linear EST 20-MAR-2001
DEFINITION EMI_33_A11_b1_A002 Embryo 1 (EM1) Sorghum bicolor cDNA, mRNA
sequence.

ACCESSION BG465076

VERSION BG465076.1 GI:13393954

KEYWORDS EST.

SOURCE Sorghum.

ORGANISM Sorghum bicolor
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
clade; Panicoideae; Andropogoneae; Sorghum.

REFERENCE 1 (bases 1 to 490)

AUTHORS Reid, S.P., Cordonnier-Pratt, M.-M., Gingle, A. and Pratt, L.H.

TITLE An EST database from Sorghum: developing embryos

JOURNAL Unpublished (2000)

COMMENT Contact: Cordonnier-Pratt MM

Laboratory for Genomics and Bioinformatics

The University of Georgia, Department of Plant Biology

Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA

Tel: 706 542 1860

Fax: 706 583 0210

Email: mmpratt@uga.edu

Sequences have been trimmed to exclude PolyA, vector and regions

below Phred quality 16. The threshold for highest quality sequence

is 20.

Seq primer: JEN REV

High quality sequence stop: 484

POLYA=No.

Location/Qualifiers

1. 490

/organism="Sorghum bicolor"

/db_xref="taxon:4558"

/clone_lib="Embryo 1 (EM1)"

/note="Organ: Embryos germinated for 24 hr; Vector:

pBluescript II from Lambda Zap II; Site_1: XhoI; Site_2:

EcoRI; The library was made from poly-A RNA in the cloning

vector lambda ZAP II. Clones to be sequenced were

prepared by mass excision."

74 a 199 c 178 g 38 t 1 others

BASE COUNT

ORIGIN

Query Match 80.9%; Score 17.8; DB 12; Length 490;
Best Local Similarity 90.5%; Pred. No. 1.8e+03;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 CCTTGGCGCTGGCGCTCAG 22

|||||
Db 345 CCTTGGCGCTGGCGCTCAG 325

RESULT 8

LOCUS AV433340 512 bp mRNA linear EST 23-AUG-2000
DEFINITION AV433340 Porphyra yezoensis TU-1 Porphyra yezoensis cDNA clone
PM027805 x 5', mRNA sequence.

ACCESSION AV433340

VERSION AV433340.1 GI:8588565

KEYWORDS EST.

SOURCE Porphyra yezoensis.

ORGANISM Porphyra yezoensis

Eukaryota; Rhodophyta; Bangiophyceae; Bangiales; Bangiaceae;

Porphyra. 1 (bases 1 to 512)

REFERENCE 1 (bases 1 to 512)

AUTHORS Nikaido, I., Asamizu, E., Nakajima, M., Nakamura, Y., Saga, N. and

Tabata, S.

TITLE Generation of 10,154 expressed sequence tags from a leafy

gametophyte of a marine red alga, Porphyra yezoensis

DNA Res. 7, 223-227 (2000)

20363100

Contact: Erika Asamizu

The First Laboratory for Plant Gene Research

Kazusa DNA Research Institute

Yana 1532-3, Kisarazu, Chiba 292-0812, Japan

Email: asamizu@kazusa.or.jp, URL: http://www.kazusa.or.jp/en/plant/.

Location/Qualifiers

1. 512

/organism="Porphyra yezoensis"

/strain="TU-1"

/db_xref="taxon:2788"

/clone_lib="PM027805"

/clone_lib="Porphyra yezoensis TU-1"

/note="Vector: pBluescriptII SK-; Site_1: EcoRI; Site_2:

XhoI"

BASE COUNT 106 a 159 c 154 g 93 t

ORIGIN

Query Match 80.9%; Score 17.8; DB 10; Length 512;

Best Local Similarity 90.5%; Pred. No. 1.8e+03;

Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 CCTTGGCGCTGGCGCTCAG 22

|||||

Db 212 CCTTGGCGCTGGCGCTCAG 232

|||||

RESULT 9

LOCUS BE254482/c 708 bp mRNA linear EST 13-JUL-2000

DEFINITION 60111575f1 NIH_MGC_16 Homo sapiens cDNA clone IMAGE:3352314 5',

mRNA sequence.

ACCESSION BE254482

VERSION BE254482.1 GI:9124910

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 708)

AUTHORS NIH-MGC http://mgc.nci.nih.gov/.

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished (1999)

COMMENT Contact: Robert Strausberg, Ph.D.

Email: cgapbs@mail.nih.gov

Tissue Procurement: ATCC
 cDNA Library Preparation: Ling Hong/Rubin Laboratory
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone Distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
 Plate: LLCW151 row: e column: 19
 High quality sequence stop: 588.
 Location/Qualifiers

FEATURES source

```

1. 708
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:3352314"
/clone_lib="NIH_MGC_16"
/tissue_type="retinoblastoma"
/lab_host="DH10B (phage-resistant)"
/note="Organ: eye; Vector: pOTB7; Site_1: XhoI; Site_2:
EcoRI; cDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adaptor: GGACGAG(G). Library constructed by Ling Hong
in the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library."
BASE COUNT      236 a 159 c 188 g 125 t
ORIGIN
Query Match      80.9%; Score 17.8; DB 10; Length 708;
Best Local Similarity 90.5%; Pred. No. 1.9e+03;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 CCTTGGCGCTGGCGGCTCTAG 22
|||||
Db 551 CCTTGGCGCTGGCGGCTCTAG 531
|||||

```

```

RESULT 10
BG699633/c
LOCUS
DEFINITION
602679365F1 NIH_MGC_95 Homo sapiens cDNA clone IMAGE:4811824 5',
mRNA sequence.
ACCESSION
BG699633
VERSION
BG699633.1 GI:13968143
KEYWORDS
EST.
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NIH-MGC http://mgi.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Straubeberg, Ph.D.
Email: cgapbs-@email.nih.gov
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki
Toshiyuki and Piero Carninci (RIKEN)
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCW10703 row: b column: 17
High quality sequence stop: 616.
Location/Qualifiers
1. 767
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4811824"
/clone_lib="NIH_MGC_95"
/tissue_type="hippocampus"
/lab_host="DH10B"
/note="Organ: brain; Vector: pBluescriptR (modified)

```

FEATURES source

```

1. 767
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4811824"
/clone_lib="NIH_MGC_95"
/tissue_type="hippocampus"
/lab_host="DH10B"
/note="Organ: brain; Vector: pBluescriptR (modified)

```

```

pBluescript KS+); Site_1: BamHI; Site_2: SalI-XhoI (gtcgag
); Oligo-dT primed using primer 5'-TTTTTTTTTTTTTTVN-3',
size-selected for average insert size 2.5 kb and
normalized to ROT 5. This is a primary library enriched
for full-length clones and constructed using the
Cap-trapper method (Carninci, in preparation). Library
constructed by M. Brownstein (NHGRI, National
Institutes of Health). Note: this is a NIH_MGC Library."
BASE COUNT      156 a 259 c 242 g 110 t
ORIGIN
Query Match      80.9%; Score 17.8; DB 12; Length 767;
Best Local Similarity 90.5%; Pred. No. 2e+03;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 CCTTGGCGCTGGCGGCTCTAG 22
|||||
Db 53 CCTTGGCGCTGGCGGCTCTAG 33
|||||

```

```

RESULT 11
AL519496/c
LOCUS
DEFINITION
AL519496 LTI_NFL004_NBC2 Homo sapiens cDNA linear EST 13-FEB-2001
prime, mRNA sequence.
ACCESSION
AL519496
VERSION
AL519496.1 GI:12782989
KEYWORDS
EST.
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Li, W.B., Gruber, C., Jesse, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished (2001)
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqreg@genoscope.cns.fr, Web : www.genoscope.cns.fr.
Location/Qualifiers
1. 856
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CS0DB004Y122"
/clone_lib="LTI_NFL004_NBC2"
/sex="male"
/tissue_type="neuroblastoma cells"
/lab_host="DH10B"
/note="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA
was primed with a NotI-oligo(dT) primer. Five prime end
enriched, double-stranded cDNA was digested with Not I and
cloned into the Not I and Eco RV sites of the pCMVSPORT 6
vector. Library was normalized. Library was constructed
by Life Technologies. Contact : Feng Liang Life
Technologies, a division of Invitrogen 9800 Medical Center
Drive Rockville, Maryland 20850, USA fax : (1) 301 610
8371 Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com"
BASE COUNT      169 a 289 c 257 g 139 t
ORIGIN
Query Match      80.9%; Score 17.8; DB 9; Length 856;
Best Local Similarity 90.5%; Pred. No. 2e+03;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 CCTTGGCGCTGGCGGCTCTAG 22
|||||
Db 26 CCTTGGCGCTGGCGGCTCTAG 6
|||||

```

RESULT 12 AL528423/c

Bouneau, L., Billault, A., Quetier, F., Saurin, W., Bernot, A. and Weissenbach, J. Characterization and repeat analysis of the compact genome of the freshwater pufferfish *Tetraodon nigroviridis*

Unpublished Genoscope.
3 (bases 1 to 936)
Submitted (12-APR-2000)
This sequence is a single read and was generated as part of a large scale clone-end sequencing project of the *Tetraodon nigroviridis* genome. For more information, please take a look at <http://www.genoscope.cns.fr/Tetraodon>.

Location/Qualifiers
1. 936
/organism="Tetraodon nigroviridis"
/db_xref="taxon:99883"
/clone_lib="214B09"
/notes="Genoscope sequence ID : COAG214CA05SP1-end : PUC-ori"

BASE COUNT 247 a 204 c 190 g 276 t 19 others
ORIGIN

Query Match 80.9%; Score 17.8; DB 17; Length 936;
Best Local Similarity 90.5%; Pred. No. 2e+03; 2; Indels 0; Gaps 0;
Matches 19; Conservative 0; Mismatches 0;

QY 1 TCCTTGGCGCTGGCGCTCA 21
|||||
Db 33 TCCTCGCGCTGGCGCCACA 53
|||||

RESULT 14
BQ684757/c 938 bp mRNA linear EST 15-JUL-2002
LOCUS AGENCOURT_8343533 NIH_MGC_110 Homo sapiens cDNA clone IMAGE:6250259
DEFINITION 5', mRNA sequence.
BQ684757
VERSION BQ684757.1 GI:21810073
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 938)
NIH-MGC <http://mgi.nci.nih.gov/>
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabs-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Plate: LLCM2393 row: a column: 12
High quality sequence stop: 710.
Location/Qualifiers
1. 938
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:6250259"
/clone_lib="NIH_MGC_110"
/lab_host="DH10B (phage-resistant)"
/note="Organ: pancreas; Vector: pOTB7; Site 1: XhoI; Site 2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University

AL528423 905 bp mRNA linear EST 13-FEB-2001
LOCUS AL528423 LTI_NFL003_NBC3 Homo sapiens cDNA clone CS0DC029YH22 5
DEFINITION prime, mRNA sequence.
ACCESSION AL528423
VERSION AL528423.1 GI:12791916
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 905)
Li, W.B., Gruber, C., Jesse, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished (2001)
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.

Location/Qualifiers
1. 905
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CS0DC029YH22"
/clone_lib="LTI_NFL003_NBC3"
/sex="male"
/tissue_type="neuroblastoma cells"
/lab_host="DH10B"
/note="Organ: brain; vector: pCMVSPORT 6; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies. Contact : Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive Rockville, Maryland 20850, USA Fax : (1) 301 610 8371 Email : liang@lifetech.com URL : <http://fulllength.invitrogen.com>"

BASE COUNT 183 a 302 c 273 g 145 t 2 others
ORIGIN

Query Match 80.9%; Score 17.8; DB 9; Length 905;
Best Local Similarity 90.5%; Pred. No. 2e+03; 2; Indels 0; Gaps 0;
Matches 19; Conservative 0; Mismatches 0;

QY 2 CCTTGGCGCTGGCGCTCAG 22
|||||
Db 35 CCTTGGCGCTGGCGCCAG 15
|||||

RESULT 13
CNS035RV 936 bp DNA linear GSS 15-MAY-2000
LOCUS Tetraodon nigroviridis genome survey sequence PUC-ori end of clone
DEFINITION 214B09 of library G from Tetraodon nigroviridis, genomic survey sequence.
AL229108
VERSION AL229108.1 GI:7898103
KEYWORDS GSS; genome survey sequence.
SOURCE Tetraodon nigroviridis.
ORGANISM Tetraodon nigroviridis.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei; Acanthomorpha; Acanthopterygii; Percormorpha; Tetraodontiformes; Tetraodontidae; Tetraodon.
1 (bases 1 to 936)
Roest-Crollius, H., Jaillon, O., Dasilva, C., Bouneau, L., Fisher, C., Bernot, A., Fizames, C., Wincker, P., Brottier, P., Quetier, F., Saurin, W. and Weissenbach, J.
Human gene number estimate provided by genome wide analysis using Tetraodon nigroviridis DNA sequence
Unpublished
2 (bases 1 to 936)
Roest-Crollius, H., Jaillon, O., Dasilva, C., Fizames, C., Fisher, C.,

of California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library."

BASE COUNT 190 a 317 c 285 g 146 t
ORIGIN

Query Match 80.9%; Score 17.8; DB 14; Length 938;
Best Local Similarity 90.5%; Pred. No. 2e+03;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 CCTTGGCGCTGGCGCTCAG 22
|||||
Db 26 CCTTGGCGCTGGCGCTCAG 6
|||||

RESULT 15
BG617183/c
LOCUS
DEFINITION BG617183 961 bp mRNA linear EST 18-APR-2001
602615581F1 NIH_MGC_76 Homo sapiens cDNA clone IMAGE:4734237 5',
mRNA sequence.
ACCESSION BG617183
VERSION BG617183
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: CLONETECH Laboratories, Inc.
cDNA Library Preparation: CLONETECH Laboratories, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LCM1597 row: a column: 22
High quality sequence stop: 211.

FEATURES
source
1..961
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4734237"
/clone_lib="NIH_MGC_76"
/lab_host="DH10B (TI phage-resistant)"
/note="Organ: liver; Vector: pDNR-LTB (Clontech); Site 1:
SfiI (ggcgctcgcc); Site 2: SfiI (ggcattatggcc); 5' and
3' adaptors were used in cloning as follows: 5' adaptor
sequence: 5'-CAGCGCATTTGCCC-3' and 3' adaptor sequence:
5'-ATTCTAGCGCGCGCGCGCATG-dt(30)BN-3' (where B = A,
C, or G and N = A, C, G, or T). Average insert size 1.85
kb (range 1.0-4.0 kb). 15/15 colonies contained inserts
by PCR. This library was enriched for full-length clones
and was constructed by Clontech Laboratories (Palo Alto,
CA). Note: this is a NIH_MGC Library."

BASE COUNT 253 a 216 c 408 g 84 t
ORIGIN

Query Match 80.9%; Score 17.8; DB 12; Length 961;
Best Local Similarity 90.5%; Pred. No. 2e+03;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 TCCTTGGCGCTGGCGCTCA 21
|||||
Db 704 TCCTTGGCGCTGGCGCTCA 684
|||||

Search completed: April 18, 2003, 07:31:53
Job time : 770.167 secs

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 18, 2003, 04:46:26 ; Search time 305.273 Seconds
(without alignments)
2192.677 Million cell updates/sec

Title: US-09-270-437D-12

Perfect score: 23
Sequence: 1 ccaactgggtggccattcagcttc 23

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:

- 1: gb_ba.*
- 2: gb_htg.*
- 3: gb_in.*
- 4: gb_om.*
- 5: gb_ov.*
- 6: gb_pat.*
- 7: gb_ph.*
- 8: gb_pl.*
- 9: gb_pr.*
- 10: gb_ro.*
- 11: gb_sts.*
- 12: gb_sy.*
- 13: gb_un.*
- 14: gb_vi.*
- 15: em_ba.*
- 16: em_fun.*
- 17: em_hum.*
- 18: em_in.*
- 19: em_mu.*
- 20: em_om.*
- 21: em_or.*
- 22: em_ov.*
- 23: em_pat.*
- 24: em_ph.*
- 25: em_pl.*
- 26: em_ro.*
- 27: em_sts.*
- 28: em_un.*
- 29: em_vi.*
- 30: em_htg_hum.*
- 31: em_htg_in.*
- 32: em_htg_other.*
- 33: em_htg_mus.*
- 34: em_htg_pln.*
- 35: em_htg_rdt.*
- 36: em_htg_mam.*
- 37: em_htg_vrt.*
- 38: em_sy.*
- 39: em_htgo_hum.*
- 40: em_htgo_mus.*
- 41: em_htgo_other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
C 1	23	100.0	1708	6	AR171864	AR171864 Sequence
C 2	23	100.0	1946	6	AR171866	AR171866 Sequence
C 3	23	100.0	2130	9	AF117106	AF117106 Homo sapi
C 4	23	100.0	2381	9	AF198254	AF198254 Homo sapi
C 5	23	100.0	65435	2	AC104587	AC104587 Homo sapi
C 6	23	100.0	147767	2	AC105030	AC105030 Homo sapi
C 7	23	100.0	154604	2	AC104974	AC104974 Homo sapi
C 8	23	100.0	159122	2	AC025556	AC025556 Homo sapi
C 9	23	100.0	168613	9	AC091133	AC091133 Homo sapi
C 10	19.8	86.1	299350	1	AP005370	AP005370 Thermosyn
C 11	19.4	84.3	54441	2	AC101033	AC101033 Mus muscu
C 12	19.4	84.3	54441	2	AC101033	AC101033 Mus muscu
C 13	19.4	84.3	191918	2	AC102692	AC102692 Mus muscu
C 14	19.4	84.3	277892	2	AC125105	AC125105 Mus muscu
C 15	18.8	81.7	1520	9	BC013150	BC013150 Homo sapi
C 16	18.8	81.7	1648	9	BC015866	BC015866 Homo sapi
C 17	18.8	81.7	2526	3	AY052019	AY052019 Drosophill
C 18	18.8	81.7	3098	9	AK091249	AK091249 Homo sapi
C 19	18.8	81.7	14668	2	AC018286	AC018286 Drosophill
C 20	18.8	81.7	149433	9	AL135901	AL135901 Human DNA
C 21	18.8	81.7	152443	3	AC009208	AC009208 Drosophill
C 22	18.8	81.7	179191	2	AL808114	AL808114 Mus muscu
C 23	18.8	81.7	181542	9	AL137019	AL137019 Human DNA
C 24	18.8	81.7	182494	3	AC009343	AC009343 Drosophill
C 25	18.8	81.7	205336	10	AC116576	AC116576 Mus muscu
C 26	18.8	81.7	226950	10	AL671011	AL671011 Mouse DNA
C 27	18.8	81.7	257909	3	AE003828	AE003828 Drosophill
C 28	18.4	80.0	131145	9	HS277661	AJ277661 Homo sapi
C 29	18.4	80.0	168097	2	AC102401	AC102401 Mus muscu
C 30	18.4	80.0	174986	9	HS277662	AJ277662 Homo sapi
C 31	18.4	80.0	213873	2	AC019200	AC019200 Homo sapi
C 32	18.4	80.0	223741	9	AC091013	AC091013 Homo sapi
C 33	18.2	79.1	10918	1	AE013239	AE013239 Methanosa
C 34	18.2	79.1	11562	1	AE009689	AE009689 Brucella
C 35	18.2	79.1	133451	9	AL158847	AL158847 Human DNA
C 36	18.2	79.1	145327	9	AP004285	AP004285 Homo sapi
C 37	18.2	79.1	152760	2	AC106146	AC106146 Rattus no
C 38	18.2	79.1	155704	9	AC073330	AC073330 Homo sapi
C 39	18.2	79.1	156461	9	AC004383	AC004383 Human Chr
C 40	18.2	79.1	168140	2	AC013379	AC013379 Homo sapi
C 41	18.2	79.1	168949	9	AC046134	AC046134 Homo sapi
C 42	18.2	79.1	175882	2	AC118532	AC118532 Rattus no
C 43	18.2	79.1	176128	2	AC124533	AC124533 Mus muscu
C 44	18.2	79.1	178388	9	AC079630	AC079630 Homo sapi
C 45	18.2	79.1	178588	2	AC118834	AC118834 Rattus no

ALIGNMENTS

RESULT 1	AR171864/c	AR171864	1708 bp	DNA	linear	PAT 17-DEC-2001
LOCUS	Sequence 5 from patent US 6297364.					
DEFINITION	Sequence 5 from patent US 6297364.					
ACCESSION	AR171864					
VERSION	AR171864.1	GI:17910814				
KEYWORDS	Unknown.					
SOURCE	Unknown.					
ORGANISM	Unclassified.					
REFERENCE	1 (bases 1 to 1708)					
AUTHORS	Chen, Y.-T., Gure, A., Tsang, S., Stockert, E., Jager, E., Alexander, K.					
TITLE	Isolated nucleic acid molecule encoding cancer associated antigen, and the antigen itself, and uses thereof					

/product="mRNA-binding protein CRDBP"
/protein id="AAF37203.1"
/db_xref="GI:7141072"

/translation="MNKLYIGNLNEVTPADLEKFAHKKISYSGQFLVKSQYAFVDC
PDEHWAMKAITFSGKVLQGRLEIEHSVPKQSRKQIRNIPPOLRWEVLSLLA
QYGVNCEQVETSETAVNTVYNSRQTRQAIKMLNGHLENHAKLVSYIPDEQIA
CGPENGRRGSGFSGPQGGPVAAGAPAKQQQVDIPLRLVPTQYVGAIGKEGATV
VPLKTLAHNFVGLRIGKEGRNLKVEQDTETKITISLQDLTYLNPRTITVKGAL
NCCRAEQIIMKKVREAYENDVAMSLQSHLTPGLNLAAGLGPASSAAVPPSPSVTG
AAYPSFNQAEQEVNVQFIPAQAVGAILGGQHIKQISRFASAKIAPPTPDSK
VRMIIITPPQAQQAQRIYKLEENFFGKEEVKLETHIRVPSAAGRVKGGKK
TVNELQNTAAEVVVPDQTPDNDQVIVKIIGHFYASQMAQRKIRDILAQVKQHQK
GQSNQAQARRK"

BASE COUNT 596 a 687 c 664 g 434 t

Query Match 100.0%; Score 23; DB 9; Length 2381;
Best Local Similarity 100.0%; Pred. No. 0.96;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CCAACTGGTGGCAATCAGCTTC 23

Db 738 CCAACTGGTGGCAATCAGCTTC 716

RESULT 5
AC104587 65435 bp DNA linear HTG 13-DEC-2001
LOCUS Homo sapiens chromosome 17 clone RP11-78C17 map 17, LOW-PASS
DEFINITION SEQUENCE SAMPLING.

ACCESSION AC104587

VERSION AC104587.1 GI:17646987

KEYWORDS HTG; HTGS PHASE0.

SOURCE Homo sapiens.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

Birren,B., Linton,L., Nusbaum,C. and Lander,E.

1 (bases 1 to 65435)

Homo sapiens chromosome 17, clone RP11-78C17

Unpublished

2 (bases 1 to 65435)

Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,
Anderson,S., Barna,N., Baatien,V., Boguslavsky,L., Boukhgalter,B.,
Brown,A., Camarata,J., Campopiano,A., Chang,J., Chazaro,B.,
Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cook,A.,
Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S.,
Farreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S.,
Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,
Hagos,B., Heaford,A., Horton,L., Hulme,W., Iliev,I., Johnson,R.,
Jones,C., Kamat,A., Karatas,A., Kells,C., LaRocque,K.,
Lamazares,R., Landers,T., Lehoczy,J., Levine,R., Liu,G.,
MacLean,C., MacDonald,P., McKernan,K., Marquis,N., Matthews,C.,
McCarthy,M., McEwan,P., Major,J., McPheters,R., Meldrim,J.,
Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C.,
Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D.,
Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V.,
Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P.,
Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupback,R.,
Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
Strauss,N., Subramanian,A., Talamas,J., Testfaye,S., Theodore,J.,
Topham,K., Travers,M., Travis,N., Triglio,J., Vassiliev,H.,
Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

Direct Submission

Submitted (13-DEC-2001) Whitehead Institute/MIT Center for Genome

Research, 300 Charles Street, Cambridge, MA 02141, USA

All repeats were identified using RepeatMasker:

Smith, A.F.A. & Green, P. (1996-1997)

http://ftp.genome.washington.edu/RN/RepeatMasker.html

----- Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L22859
Center clone name: 78_C_17

* NOTE: This record contains 84 individual
* sequencing reads that have not been assembled into
* contigs. Runs of N are used to separate the reads
* and the order in which they appear is completely
* arbitrary. Low-pass sequence sampling is useful for
* identifying clones that may be gene-rich and allows
* overlap relationships among clones to be deduced.
* However, it should not be assumed that this clone
* will be sequenced to completion. In the event that
* the record is updated, the accession number will
* be preserved.

1 679: contig of 679 bp in length
680 779: gap of 100 bp
780 1495: contig of 716 bp in length
1496 1595: gap of 100 bp
1596 2279: contig of 684 bp in length
2280 2379: gap of 100 bp
2380 3058: contig of 679 bp in length
3059 3158: gap of 100 bp
3159 3793: contig of 635 bp in length
3794 3893: gap of 100 bp
3894 4572: contig of 679 bp in length
4573 4672: gap of 100 bp
4673 5340: contig of 668 bp in length
5341 5440: gap of 100 bp
5441 6128: contig of 688 bp in length
6129 6228: gap of 100 bp
6229 6909: contig of 681 bp in length
6910 7009: gap of 100 bp
7010 7726: contig of 717 bp in length
7727 7836: gap of 100 bp
7837 8507: contig of 681 bp in length
8508 8607: gap of 100 bp
8608 9282: contig of 675 bp in length
9283 9382: gap of 100 bp
9383 10055: contig of 673 bp in length
10056 10155: gap of 100 bp
10156 10825: contig of 670 bp in length
10826 10925: gap of 100 bp
10926 11616: contig of 691 bp in length
11617 11716: gap of 100 bp
11717 12414: contig of 698 bp in length
12415 12514: gap of 100 bp
12515 13177: contig of 663 bp in length
13178 13277: gap of 100 bp
13278 13961: contig of 684 bp in length
13962 14061: gap of 100 bp
14062 14729: contig of 668 bp in length
14730 14829: gap of 100 bp
14830 15534: contig of 705 bp in length
15535 15634: gap of 100 bp
15635 16337: contig of 703 bp in length
16338 16437: gap of 100 bp
16438 17126: contig of 689 bp in length
17127 17226: gap of 100 bp
17227 17897: contig of 671 bp in length
17898 17997: gap of 100 bp
17998 18654: contig of 657 bp in length
18655 18754: gap of 100 bp
18755 19438: contig of 684 bp in length
19439 19538: gap of 100 bp
19539 20216: contig of 678 bp in length
20217 20316: gap of 100 bp
20317 21000: contig of 684 bp in length
21001 21100: gap of 100 bp
21101 21777: contig of 677 bp in length

Mon Apr 21 10:30:06 2003

* 21778 21877: gap of 100 bp
 * 21878 22550: contig of 673 bp in length
 * 22551 22650: gap of 100 bp
 * 22651 23305: contig of 655 bp in length
 * 23306 23405: gap of 100 bp
 * 23406 24108: contig of 703 bp in length
 * 24109 24208: gap of 100 bp
 * 24209 24911: contig of 703 bp in length
 * 24912 25011: gap of 100 bp
 * 25012 25697: contig of 686 bp in length
 * 25698 25797: gap of 100 bp
 * 25798 26455: contig of 658 bp in length
 * 26456 26555: gap of 100 bp
 * 26556 27238: contig of 683 bp in length
 * 27239 27338: gap of 100 bp
 * 27339 28005: contig of 667 bp in length
 * 28006 28105: gap of 100 bp
 * 28106 28794: contig of 689 bp in length
 * 28795 28894: gap of 100 bp
 * 28895 29554: contig of 660 bp in length
 * 29555 29654: gap of 100 bp
 * 29655 30351: contig of 697 bp in length
 * 30352 30451: gap of 100 bp
 * 30452 31136: contig of 685 bp in length
 * 31137 31236: gap of 100 bp
 * 31237 31922: contig of 686 bp in length
 * 31923 32022: gap of 100 bp
 * 32023 32693: contig of 671 bp in length
 * 32694 32793: gap of 100 bp
 * 32794 33492: contig of 699 bp in length
 * 33493 33592: gap of 100 bp
 * 33593 34301: contig of 709 bp in length
 * 34302 34401: gap of 100 bp
 * 34402 35025: contig of 624 bp in length
 * 35026 35125: gap of 100 bp
 * 35126 35799: contig of 674 bp in length
 * 35800 35899: gap of 100 bp
 * 35900 36548: contig of 649 bp in length
 * 36549 36648: gap of 100 bp
 * 36649 37327: contig of 679 bp in length
 * 37328 37427: gap of 100 bp
 * 37428 38120: contig of 693 bp in length
 * 38121 38220: gap of 100 bp
 * 38221 38910: contig of 690 bp in length
 * 38911 39010: gap of 100 bp
 * 39011 39681: contig of 671 bp in length
 * 39682 39781: gap of 100 bp
 * 39782 40458: contig of 677 bp in length
 * 40459 40558: gap of 100 bp
 * 40559 41216: contig of 658 bp in length
 * 41217 41316: gap of 100 bp
 * 41317 42019: contig of 703 bp in length
 * 42020 42119: gap of 100 bp
 * 42120 42812: contig of 693 bp in length
 * 42813 42912: gap of 100 bp
 * 42913 43607: contig of 695 bp in length
 * 43608 43707: gap of 100 bp
 * 43708 44395: contig of 688 bp in length
 * 44396 44495: gap of 100 bp
 * 44496 45151: contig of 656 bp in length
 * 45152 45251: gap of 100 bp
 * 45252 45927: contig of 676 bp in length
 * 45928 46027: gap of 100 bp
 * 46028 46715: contig of 688 bp in length
 * 46716 46815: gap of 100 bp
 * 46816 47517: contig of 702 bp in length
 * 47518 47617: gap of 100 bp
 * 47618 48286: contig of 669 bp in length
 * 48287 48386: gap of 100 bp
 * 48387 49073: contig of 687 bp in length
 * 49074 49173: gap of 100 bp
 * 49174 49848: contig of 675 bp in length
 * 49849 49948: gap of 100 bp

* 49949 50613: contig of 665 bp in length
 * 50614 50713: gap of 100 bp
 * 50714 51405: contig of 692 bp in length
 * 51406 51505: gap of 100 bp
 * 51506 52203: contig of 698 bp in length
 * 52204 52303: gap of 100 bp
 * 52304 52997: contig of 694 bp in length
 * 52998 53097: gap of 100 bp
 * 53098 53779: contig of 682 bp in length

Query Match 100.0%; Score 23; DB 2; Length 65435;
 Best Local Similarity 100.0%; Pred. No. 1.3; Indels 0; Gaps 0;
 Matches 23; Conservative 0; Mismatches 0

QY 1 CCAACTGGTGGCCATTCAGCTTC 23
 |||||
 Db 982 CCAACTGGTGGCCATTCAGCTTC 1004
 |||||

RESULT 6

AC105030

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

AC105030 147767 bp DNA linear HTG 06-AUG-2002
 Homo sapiens chromosome 17 clone CTD-2244F11 map 17, *** SEQUENCING
 IN PROGRESS ***, 3 ordered pieces.
 AC105030
 HTG: HTGS_PHASE2; HTGS_FULLTOP; HTGS_ACTIVEFIN.
 human.
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 Birren,B., Nusbaum,C. and Lander,B.
 1 (bases 1 to 147767)
 Homo sapiens chromosome 17, clone CTD-2244F11
 Unpublished
 2 (bases 1 to 147767)
 Birren,B., Linton,L., Nusbaum,C., Lander,B., Allen,N.,
 Anderson,S., Barna,N., Bastien,V., Boguslavskiy,L., Boukhgalter,B.,
 Brown,A., Camarata,J., Campopiano,A., Chang,J., Chazaro,B.,
 Choepel,Y., Collangelo,M., Collins,S., Collymore,A., Cook,A.,
 Cooke,P., DeArelano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S.,
 Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S.,
 Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,
 Hagos,B., Heaford,A., Horton,L., Hulme,W., Iliev,I., Johnson,R.,
 Jones,C., Kamat,A., Karatas,A., Kells,C., Laroque,K.,
 Lamazares,R., Landers,T., Lehoczy,J., Levine,R., Liu,G.,
 MacLean,C., Macdonald,P., Major,J., Marquis,N., Matthews,C.,
 McCarthy,M., McEwan,P., McKernan,K., McPheeters,R., Meldrim,J.,
 Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C.,
 Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., Pollara,V.,
 Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V.,
 Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P.,
 Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schnupack,R.,
 Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
 Strauss,N., Subramanian,A., Talamas,J., Testave,S., Theodore,J.,
 Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H.,
 Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
 Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

TITLE

JOURNAL

REFERENCE

AUTHORS

Direct Submission
 Submitted (22-DEC-2001) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA
 3 (bases 1 to 147767)
 Birren,B., Nusbaum,C., Lander,B., Allen,N., Anderson,S.,
 Barna,N., Bastien,V., Bloom,T., Boguslavskiy,L., Boukhgalter,B.,
 Camarata,J., Chang,J., Chazaro,B., Choepel,Y., Collymore,A.,
 Cook,A., Cooke,P., DeArelano,K., Dewar,K., Diaz,J.S., Dodge,S.,
 Faro,S., Ferreira,P., FitzGerald,M., Gage,D., Galagan,J.,
 Gardyna,S., Gord,S., Graham,L., Grand-Pierre,N., Hagos,B.,
 Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A.,
 Karatas,A., Kells,C., Landers,T., Levine,R., Lindblad-Toh,K.,
 Liu,G., MacLean,C., Macdonald,P., Major,J., Matthews,C.,
 McCarthy,M., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V.,
 Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., Norman,C.H.,
 O'Connor,T., O'Donnell,P., Pollara,V.,
 Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V.,
 Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P.,
 Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schnupack,R.,
 Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
 Strauss,N., Subramanian,A., Talamas,J., Testave,S., Theodore,J.,
 Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H.,
 Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
 Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

O'Connor, T., O'Donnell, P., O'Neil, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Raymond, C., Retta, R., Rise, C., Rogov, P., Roman, J., Roy, A., Schauer, S., Schupback, R., Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

Direct Submission
Submitted (06-AUG-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
On Aug 6, 2002 this sequence version replaced gi:21426162.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html

----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L23011
Center clone name: 2244_F_11

* NOTE: This is a 'working draft' sequence. It currently
* consists of 3 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* provided by the submitter.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
* 1 64310: contig of 64310 bp in length
* 64311 64410: gap of 100 bp
* 64411 86773: contig of 22363 bp in length
* 86774 86873: gap of 100 bp
* 86874 147767: contig of 60894 bp in length.
Location/Qualifiers
1. 147767
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="17"
/map="17"
/clone="CTD-2244F11"
/clone_lib="CITDI Human BAC"
BASE COUNT 41792 a 34009 c 33361 g 38110 t 495 others

Query Match 100.0%; Score 23; DB 2; Length 147767;
Best Local Similarity 100.0%; Pred. No. 1.4;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 CCACTGGTGGCCATTCAGCTTC 23
|||||
Db 106124 CCACTGGTGGCCATTCAGCTTC 106146

RESULT 7
AC104974/c
LOCUS AC104974 154604 bp DNA linear HTG 13-FEB-2002
DEFINITION Homo sapiens chromosome 17 clone RP11-145L16 map 17, WORKING DRAFT
SEQUENCE, 23 ordered pieces.
AC104974
VERSION AC104974.3 GI:18653683
KEYWORDS HTG; HTGS_PHASE2; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 154604)
Birren, B., Linton, L., Nusbaum, C. and Lander, E.
REFERENCE
AUTHORS
TITLE
JOURNAL
Unpublished

REFERENCE
AUTHORS

2 (bases 1 to 154604)
Birren, B., Linton, L., Nusbaum, C., Lander, E., Ali, A., Allen, N., Anderson, S., Barna, N., Bastien, V., Boguslavskiy, L., Boukhgalter, B., Brown, A., Camarata, J., Campopiano, A., Chang, J., Chazaro, B., Choquel, Y., Colangelo, M., Collins, S., Collumore, A., Cook, A., Cooke, P., DeArillano, K., Dewar, K., Diaz, J.S., Dodge, S., Faro, S., Ferreira, P., FitzHugh, W., Gage, D., Galagan, J., Gardyna, S., Ginde, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N., Hagos, B., Heaford, A., Horton, L., Hulme, W., Iliiev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kells, C., LaRocque, K., Lamazares, R., Landers, T., Lehoczy, J., Levine, R., Liu, G., MacLean, C., Macdonald, P., Major, J., Marquis, N., Mathews, C., McCarthy, M., McEwan, P., McKernan, K., McPheeters, R., Meldrum, J., Meneus, L., Mihova, T., Mienga, V., Murphy, T., Naylor, J., Nguyen, C., Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neil, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V., Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P., Roman, J., Roy, A., Santos, R., Schauer, S., Schupback, R., Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N., Straus, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M., Travis, N., Trigilio, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

Direct Submission
Submitted (22-DEC-2001) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
On Feb 13, 2002 this sequence version replaced gi:18464184.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html

----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L22871
Center clone name: 145_L16

----- Summary Statistics
Sequencing vector: Plasmid; n/a; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 148031 bases at least Q40
Consensus quality: 150405 bases at least Q30
Consensus quality: 151462 bases at least Q20
Insert size: 146000; agarose-fp
Insert size: 152404; sum-of-contigs
Quality coverage: 9.3 in Q20 bases; agarose-fp
Quality coverage: 8.9 in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently
* consists of 23 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* provided by the submitter.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
* 1 1709: contig of 1709 bp in length
* 1710 1809: gap of 100 bp
* 1810 1999: contig of 190 bp in length
* 2000 2099: gap of 100 bp
* 2100 3127: contig of 1028 bp in length
* 3128 3227: gap of 100 bp
* 3228 4561: contig of 1334 bp in length
* 4562 4661: gap of 100 bp
* 4662 5811: contig of 1150 bp in length
* 5812 5911: gap of 100 bp
* 5912 7151: contig of 1240 bp in length
* 7152 7251: gap of 100 bp
* 7252 8261: contig of 1010 bp in length
* 8262 8361: gap of 100 bp

TITLE
JOURNAL

COMMENT

* 8362 9739: contig of 1378 bp in length
* 9740 9839: gap of 100 bp
* 9840 12105: contig of 2266 bp in length
* 12106 12205: gap of 100 bp
* 12206 35760: contig of 23555 bp in length
* 35761 35860: gap of 100 bp
* 35861 38000: contig of 2140 bp in length
* 38001 38100: gap of 100 bp
* 38101 41369: contig of 3269 bp in length
* 41370 41469: gap of 100 bp
* 41470 44094: contig of 2625 bp in length
* 44095 44194: gap of 100 bp
* 44195 47400: contig of 3206 bp in length
* 47401 47500: gap of 100 bp
* 47501 53741: contig of 6241 bp in length
* 53742 53841: gap of 100 bp
* 53842 60833: contig of 6992 bp in length
* 60834 60933: gap of 100 bp
* 60934 72421: contig of 11488 bp in length
* 72422 72521: gap of 100 bp
* 72522 84887: contig of 12366 bp in length
* 84888 84987: gap of 100 bp
* 84988 95600: contig of 14613 bp in length
* 95601 99700: gap of 100 bp
* 99701 115353: contig of 15653 bp in length
* 115354 115453: gap of 100 bp
* 115454 131255: contig of 15802 bp in length
* 131256 131355: gap of 100 bp
* 131356 152505: contig of 21150 bp in length
* 152506 152605: gap of 100 bp
* 152606 154604: contig of 1999 bp in length.

FEATURES

source

Location/Qualifiers
1. .154604
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="17"
/map="17"
/clone="RP11-145L16"
/clone_lib="RPC1-11 Human Male BAC"

misc_feature

1. .1709
/note="assembly_fragment"
clone end:SP6
vector side:left

misc_feature

1810. .1999
/note="assembly_fragment"
2100. .3127
/note="assembly_fragment"
3228. .4561
/note="assembly_fragment"

misc_feature

4662. .5811
/note="assembly_fragment"
5912. .7151
/note="assembly_fragment"

misc_feature

7252. .8261
/note="assembly_fragment"
8362. .9739
/note="assembly_fragment"

misc_feature

9840. .12105
/note="assembly_fragment"
12206. .35760
/note="assembly_fragment"

misc_feature

35861. .38000
/note="assembly_fragment"
38101. .41369
/note="assembly_fragment"

misc_feature

41470. .44094
/note="assembly_fragment"
44195. .47400
/note="assembly_fragment"

misc_feature

47501. .53741
/note="assembly_fragment"
53842. .60833
/note="assembly_fragment"

misc_feature 60934. .72421
/note="assembly_fragment"
misc_feature 72522. .84887
/note="assembly_fragment"
misc_feature 84988. .99600
/note="assembly_fragment"
misc_feature 99701. .115353
/note="assembly_fragment"
misc_feature 115454. .131255
/note="assembly_fragment"
misc_feature 131356. .152505
/note="assembly_fragment"
misc_feature 152606. .154604
/note="assembly_fragment"
clone end:T7
vector side:right

BASE COUNT 42238 a 34832 c 34671 g 40661 t 2202 others
ORIGIN

Query Match 100.0%; Score 23; DB 2; Length 154604;
Best Local Similarity 100.0%; Pred.No.1.4;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCAACTGGTGGCCATTTCAGCTTC 23
|||||
DB 133577 CCAACTGGTGGCCATTTCAGCTTC 133555

RESULT 8

AC025556

LOCUS

DEFINITION

SEQUENCE, 17 unordered pieces.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

On Sep 1, 2000 this sequence version replaced gi:8570253.

Center: Washington University Genome Sequencing Center

Center code: WUGSC

Web site: <http://genome.wustl.edu/gsc/index.shtml>

Project Information

Center project name: H_NH0501C14

Summary Statistics

Sequencing vector: M13; 96%

Chemistry: Dye-terminator ET; 96% of reads

Chemistry: Dye-terminator Big Dye; 4% of reads

Assembly program: Phrap; version 0.990319

Consensus quality: 148686 bases at least Q40

Consensus quality: 152251 bases at least Q30

Consensus quality: 154277 bases at least Q20

Insert size: 181000; agarose-fp

Insert size: 157522; sum-of-contigs

Quality coverage: 3.86 in Q20 bases; agarose-fp

Quality coverage: 4.48 in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently

AC025556 159122 bp DNA linear HTG 01-SEP-2000
Homo sapiens chromosome 17 clone RP11-501C14, WORKING DRAFT
SEQUENCE, 17 unordered pieces.

AC025556
AC025556.4 GI:9958273
HTG; HTGS_PHASE1; HTGS_DRAFT.
Homo sapiens.

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Waterston,R.H.
The sequence of Homo sapiens clone

1 (bases 1 to 159122)

Unpublished

Waterston,R.H.

2 (bases 1 to 159122)

Direct Submission

Submitted (10-MAR-2000) Genome Sequencing Center, Washington

University School of Medicine, 4444 Forest Park Parkway, St. Louis,

MO 63108, USA

On Sep 1, 2000 this sequence version replaced gi:8570253.

Center: Washington University Genome Sequencing Center

Center code: WUGSC

Web site: <http://genome.wustl.edu/gsc/index.shtml>

Project Information

Center project name: H_NH0501C14

Summary Statistics

Sequencing vector: M13; 96%

Chemistry: Dye-terminator ET; 96% of reads

Chemistry: Dye-terminator Big Dye; 4% of reads

Assembly program: Phrap; version 0.990319

Consensus quality: 148686 bases at least Q40

Consensus quality: 152251 bases at least Q30

Consensus quality: 154277 bases at least Q20

Insert size: 181000; agarose-fp

Insert size: 157522; sum-of-contigs

Quality coverage: 3.86 in Q20 bases; agarose-fp

Quality coverage: 4.48 in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently

* consists of 17 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

```

1
7635: contig of 7635 bp in length
7736: gap of unknown length
15480: contig of 7745 bp in length
15581: gap of unknown length
25089: contig of 9509 bp in length
25190: gap of unknown length
35006: contig of 9817 bp in length
35106: gap of unknown length
35107: contig of 9403 bp in length
44509: gap of unknown length
44610: contig of 11992 bp in length
56601: gap of unknown length
56701: contig of 7835 bp in length
64536: gap of unknown length
64637: contig of 15305 bp in length
79941: contig of 15705 bp in length
80042: gap of unknown length
95746: contig of 17975 bp in length
95847: gap of unknown length
125322: contig of 29686 bp in length
125333: contig of 1179 bp in length
126811: gap of unknown length
126912: contig of 2235 bp in length
129146: gap of unknown length
129247: contig of 3446 bp in length
132693: contig of 3122 bp in length
132793: gap of unknown length
135914: contig of 7546 bp in length
136015: gap of unknown length
143660: contig of 7106 bp in length
143661: gap of unknown length
150767: contig of 8256 bp in length.
150867: Location/Qualifiers
1. .159122
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="17"
/clones="RP11-501C14"
1. .7635
/notes="assembly_name:Contig10"
7736. .15480
/notes="assembly_name:Contig11"
15581. .25089
/notes="assembly_name:Contig12"
25190. .35006
/notes="assembly_name:Contig13"
35107. .44509
/notes="assembly_name:Contig14"
44610. .56601
/notes="assembly_name:Contig15"
56702. .64536
/notes="assembly_name:Contig16"
64637. .79941
/notes="assembly_name:Contig17"
80042. .95746
/notes="assembly_name:Contig18"
95847. .125322
/notes="assembly_name:Contig19"
125633. .126811
/notes="assembly_name:Contig3"
126912. .129146
/notes="assembly_name:Contig4"
129247. .132692
/notes="assembly_name:Contig5"
132793. .135914

```

FEATURES

```

source
misc_feature
/notes="assembly_name:Contigs"
136015..143560
/notes="assembly_name:Contig7"
143661..150766
/notes="assembly_name:Contig8"
150867..159122
/notes="assembly_name:Contig9"
42336 a 37513 c 36722 g 40933 t 1618 others
ORIGIN

Query Match
Best Local Similarity 100.0%; Score 23; DB 2; Length 159122;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CCAACTGGTGGCCATTGAGCTTC 23
|||||
Db 67199 CCAACTGGTGGCCATTGAGCTTC 67221

RESULT 9
AC091133
LOCUS
DEFINITION Homo sapiens chromosome 17, clone RP11-501C14, complete sequence.
AC091133
VERSION AC091133.11 GI:21637524
KEYWORDS HTG.
SOURCE
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 (bases 1 to 168613)
AUTHORS Birren,B., Nusbaum,C. and Lander,E.
TITLE Homo sapiens chromosome 17, clone RP11-501C14
JOURNAL Unpublished
AUTHORS
2 (bases 1 to 168613)
AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,S.,
Barna,N., Bastien,V., Boguslavskiy,L., Bouckgalter,B., Brown,A.,
Camarata,J., Campopiano,A., Chang,J., Choepel,Y., Colangelo,M.,
Collins,S., Collymore,A., Cooke,P., Dearellano,K., Dewar,K.,
Diaz,J.S., Dodge,S., Fero,S., Ferreira,P., FitzHugh,W., Gage,D.,
Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L.,
Grand-Pierre,N., Hagos,B., Heaford,A., Horton,L., Hulme,W.,
Iliev,I., Johnson,R., Jones,C., Karatas,A., Larocque,K.,
Lamazares,R., Landers,T., Lehoczy,J., Levine,R., Liu,G.,
Maclean,C., Macdonald,P., Marquis,N., Matthews,C., McCarthy,M.,
McEwan,P., McKernan,K., McPheeters,R., Meldrum,J., Meneus,L.,
Mihova,T., Mienga,V., Murphy,T., Naylor,J., Nguyen,C., Norbu,C.,
Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J.,
Peterson,K., Phunkhang,P., Pierre,N., Pollara,V., Raymond,C.,
Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P., Roman,J.,
Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupback,R., Seaman,S.,
Severy,P., Sougnez,C., Spencer,B., Stange-Thomann,N.,
Stojanovic,N., Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S.,
Theodorovic,J., Travers,M., Travis,N., Trigilio,J., Vassiliev,H.,
Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
Direct Submission
Submitted (01-APR-2001) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 168613)
AUTHORS
Anderson,S., Barna,N., Bastien,V., Bloom,T., Boguslavskiy,L.,
Bouckgalter,B., Brown,A., Camarata,J., Campopiano,A., Chang,J.,
Chazaro,B., Choepel,Y., Colangelo,M., Collins,S., Collymore,A.,
Cook,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S.,
Fero,S., Ferreira,P., FitzGerald,M., Dewar,K., FitzHugh,W., Gage,D.,
Galagan,J., Gardyna,S., Ginde,S., Gord,S., Goyette,M., Graham,L.,
Grand-Pierre,N., Hagos,B., Horton,L., Hulme,W., Iliev,I.,
Johnson,R., Jones,C., Kamat,A., Karatas,A., Kells,C., Larocque,K.,
Lamazares,R., Landers,T., Lehoczy,J., Levine,R., Lindblad-Toh,K.,
Liu,G., Maclean,C., Macdonald,P., Major,J., Marquis,N.,
Matthews,C., McCarthy,M., McEwan,P., McKernan,K., Meldrum,J.,
Meneus,L., Mihova,T., Mienga,V., Murphy,T., Naylor,J., Nguyen,C.,

```

[illegible]

TITLE JOURNAL REFERENCE AUTHORS	TITLE JOURNAL COMMENT

```

----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence.submissions@genome.wi.mit.edu
----- Project Information
Center project name: L12028
Center clone name: 501_C_14
-----
Location/Qualifiers
1. .168613
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="17"
/map="17"
/clone="RP11-501C14"
/clone_lib="RPC1-11 Human Male BAC"
119. .201
/rpt_family="(CA)n"
repeat_region
255. .519
/rpt_family="AluDb"
repeat_region
456. .460
/note="<30 quality SNGL region"
unsure
484. .488
/note="<30 qual SNGL region"
unsure
1041. .1349
/rpt_family="AluX"
repeat_region
1271. .1278
/note="<30 quality SNGL region"
unsure
1293. .1296
/note="<30 quality SNGL region"
unsure
1394. .1503
/note="single clone coverage"
unsure
1569. .1573
/note="<30 quality SNGL region"
unsure

```


QY 1 CCAACTGGTGGCCATTACGCTTC 23
DB 106850 CCAACTGGTGGCCATTACGCTTC 106828

RESULT 11
AC101033
LOCUS Mus musculus clone RP23-8514, DNA linear HTG 23-NOV-2001
DEFINITION Mus musculus clone RP23-8514, LOW-PASS SEQUENCE SAMPLING.
AC101033
ACCESSION AC101033.1 GI:17059807
VERSION HTG; HTGS PHASE0.
KEYWORDS Mus musculus.
SOURCE Mus musculus.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
Birren, B., Linton, L., Nusbaum, C. and Lander, E.
Mus musculus, clone RP23-8514
Unpublished
2 (bases 1 to 54441)
Birren, B., Linton, L., Nusbaum, C., Lander, E., Ali, A., Allen, N.,
Anderson, S., Barna, N., Bastien, V., Boguslavsky, L., Boukhgalter, B.,
Brown, A., Camarata, J., Campopiano, A., Chang, J., Chazaro, B.,
Choepe, Y., Colangelo, M., Collins, S., Collymore, A., Cook, A.,
Cooke, P., DeArellano, K., Dewar, K., Diaz, J., Dodge, S., Faro, S.,
Ferreira, P., Fitzhugh, W., Gage, D., Galagan, J., Gardyna, S.,
Ginde, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N.,
Hagos, B., Heaford, A., Horton, L., Hulme, W., Iliev, I., Johnson, R.,
Jones, C., Kamat, A., Karatas, A., Kells, C., LaRocque, K.,
Lamazares, R., Landers, T., Lehotzky, J., Levine, R., Liu, G.,
MacClean, C., MacDonald, P., Major, J., Marquis, N., Matthews, C.,
McCarthy, M., McEwan, P., McKernan, K., McPheeters, R., Meldrum, J.,
Meneus, L., Mihova, T., Mienga, V., Murphy, T., Naylor, J., Nguyen, C.,
Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neil, D.,
Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V.,
Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P.,
Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schupbach, R.,
Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N.,
Strauss, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J.,
Topham, K., Travers, M., Travis, N., Trigglio, J., Vassiliev, H.,
Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J., Young, G.,
Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

Direct Submission
Submitted (23-NOV-2001) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L14840
Center clone name: 85_1_4

* NOTE: This record contains 68 individual
* sequencing reads that have not been assembled into
* contigs. Runs of N are used to separate the reads
* and the order in which they appear is completely
* arbitrary. Low-pass sequence sampling is useful for
* identifying clones that may be gene-rich and allows
* overlap relationships among clones to be deduced.
* However, it should not be assumed that this clone
* will be sequenced to completion. In the event that
* the record is updated, the accession number will
* be preserved.
* 1 689: contig of 689 bp in length
* 690 789: gap of 100 bp
* 790 1570: contig of 781 bp in length
* 1571 1670: gap of 100 bp

TITLE
JOURNAL
COMMENT
/gene="t1r0320"
4652..7339
/gene="t1r0320"
/note="ORF ID:t1r0320
unknown protein"
/codon_start=1
/transl_table=11
/protein_id="BAC07872.1"
/db_xref="GI:22294041"
/translation="MSLNQSKTKTAQGTWEAPNLVYRKQQAARDIAQCFKGQWQOFSE
FIDCLEQTHCAQAGFELQCKADIAALHSLEALQALNARDLSPILLATLTGTT
SGKTLNFCAGAEPLVAVSEMSAGTIVLEGDVPSLTIAATPCALWECGTWENISA
ATYDKLEAMQYLDTRNPSLAPKYNATYPIQFLRQHGLPQTRVQLLOL
PGYVGRNEMAGIIFECORQALCLVNTASATDAKTEQLLEEVIVQVNLGSPAR
MLFVNLIDVDFQKDPQSEQRQFVKVQIKDKLEKPEHEAETLTKVLKLSAR
AALYSQIQSVGEERAKICERVDQHQVILGKLELPRNPAKWTDDCORVAESL
WENYQCELFELVEQVQHFPELVIPQIVADFNAGVANKVLETCOSVHALCNSIE
KYQCEBLETGVQLSEKLRAGENLKVFFEMKELITSSHDPVQVLEBKLEFIE
FIGNKEQTECEPEVDNDDEWSDKIRQWRDKIQQWHSQDDDEMDKDLDEELPLI
YTWRFVSIDKLQSVYLSIQDGKVDFAADNLKQPPDPQKVVQLEBKLEFIE
FINAIDCNADVAECMTIROEPQDKTIHFQSIQDIQKLSVNVLVKSVLARE
QGRYDLSLLFNHLEYLQESQRIAPELGQFPMSTLSQVQKELNMGISLEKAFSP
EAKETATYTERKTRPMLLWLFETITVTVQKSGLELVSPMAKLCEDWSAAKNR
VYELWPAFLDVFQTNLLTEQANEYQQAVALERYRSKLSCEHQALSIDYETRKQFWE
LHMTSVAELQEPSPWDKIKQSENAGTVSVELNAETSSV"
/complement(8059..9222)
/gene="t1l0321"
/complement(8059..9222)
/gene="t1l0321"
/note="ORF ID:t1l0321"
/codon_start=1
/transl_table=11
/product="lipid A disaccharide synthase"
/protein_id="BAC07873.1"
/db_xref="GI:22294042"
/translation="MAHFLFSTGVSGDLQALLVKALVRLAAERGMPLIEISALGGDR
MAAGAKVFLNGISGVGLLEALPLIKPTIALQLKARRVLOQHPDLVLYIDYIGN
VAMQGFTRHFAPIVYITAPQEWWSHGKTKTQOIVALSRLLAIFPEASVYRRHG
ANVVGHGLFDRIAAAPSRQVQSGIAADLALALPLSKQISLILPLIGAA
TNIKAYQPAERFPLSLQVQPAIEAVLQVPICVTLAEDSLQVLAADLAIAKSGT
VNLEFALNVQVVIYRVHPLSLYQVFLKFNLFQVSPNLLVGRVREIFVQLQDRAT
IDNTAFAALLDHPKELKMAQVGAEMRAMGTAGVVDRAATEILNLIPQNG"
/complement(9225..10364)
/gene="t1l0322"
/complement(9225..10364)
/gene="t1l0322"
/note="ORF ID:t1l0322
hypothetical protein"
/codon_start=1
/transl_table=11
/protein_id="BAC07874.1"
/db_xref="GI:22294043"
/translation="WLSKGFVELVTGPTGTEIVGLSDRIYVKDLPFGVREDSNVRF
TTPVLYLQALCDLIRPFRFLRAYLQSLGDLTVPGTSLGDTQRFYRSDPNVYH
TVTEQYGRVVTASVHINIGLREPEALMRACRLVLEAPFLFALSAPRNGDQVTH
YHSTWAIPEKTPPOVPLFTSHAFTIEWTEAQOQGTQWNVHLSAVRPNDRRVID
LNRLERICDLVTDRITALLITALEARLLQDLPDPLDPLRWGDRETLAQADENQ
LAAKRSLEAVLTHWRDRRLQITAAAMTAEIYEVWPIAKAQGFSFLVPIKILRQGT
AQQLWYAAGQTIPEMAQVQEMANEQEFADQLCPPLATVQ"
/complement(10533..11003)
/gene="t1r0323"
/complement(10533..11003)
/gene="t1r0323"
/note="ORF ID:t1r0323
unknown protein"
/codon_start=1
/transl_table=11
/protein_id="BAC07875.1"
/db_xref="GI:22294044"
/translation="WLSKGFVELVTGPTGTEIVGLSDRIYVKDLPFGVREDSNVRF
TTPVLYLQALCDLIRPFRFLRAYLQSLGDLTVPGTSLGDTQRFYRSDPNVYH
TVTEQYGRVVTASVHINIGLREPEALMRACRLVLEAPFLFALSAPRNGDQVTH
YHSTWAIPEKTPPOVPLFTSHAFTIEWTEAQOQGTQWNVHLSAVRPNDRRVID
LNRLERICDLVTDRITALLITALEARLLQDLPDPLDPLRWGDRETLAQADENQ
LAAKRSLEAVLTHWRDRRLQITAAAMTAEIYEVWPIAKAQGFSFLVPIKILRQGT
AQQLWYAAGQTIPEMAQVQEMANEQEFADQLCPPLATVQ"
/complement(11003..11003)
/gene="t1r0323"
/complement(11003..11003)
/gene="t1r0323"
/note="ORF ID:t1r0323
unknown protein"
/codon_start=1
/transl_table=11
/protein_id="BAC07875.1"

gene
CDS
Query Match 86.1%; Score 19.8; DB 1; Length 299350;
Best Local Similarity 91.3%; Pred. No. 64;
Matches 21; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
* 1671 2370: contig of 700 bp in length
* 2371 2470: gap of 100 bp
* 2471 3175: contig of 705 bp in length
* 3176 3275: gap of 100 bp
* 3276 3943: contig of 668 bp in length
* 3944 4043: gap of 100 bp
* 4044 4742: contig of 699 bp in length
* 4743 4842: gap of 100 bp
* 4843 5543: contig of 701 bp in length
* 5544 5643: gap of 100 bp
* 5644 6360: contig of 717 bp in length
* 6361 6460: gap of 100 bp
* 6461 7157: contig of 697 bp in length
* 7158 7257: gap of 100 bp
* 7258 7960: contig of 703 bp in length
* 7961 8060: gap of 100 bp
* 8061 8760: contig of 700 bp in length
* 8761 8860: gap of 100 bp
* 8861 9566: contig of 706 bp in length
* 9567 10361: contig of 695 bp in length
* 10362 10461: gap of 100 bp
* 10462 11153: contig of 692 bp in length
* 11154 11253: gap of 100 bp
* 11254 11964: contig of 711 bp in length
* 11965 12064: gap of 100 bp
* 12065 12717: contig of 653 bp in length
* 12718 12817: gap of 100 bp
* 12818 13534: contig of 717 bp in length
* 13535 13634: gap of 100 bp
* 13635 14333: contig of 699 bp in length
* 14334 14433: gap of 100 bp
* 14434 15137: contig of 704 bp in length
* 15138 15237: gap of 100 bp
* 15238 15949: contig of 712 bp in length
* 15950 16049: gap of 100 bp
* 16050 16760: contig of 711 bp in length
* 16761 16860: gap of 100 bp
* 16861 17570: contig of 710 bp in length
* 17571 17670: gap of 100 bp
* 17671 18380: contig of 710 bp in length
* 18381 18480: gap of 100 bp
* 18481 19183: contig of 703 bp in length
* 19184 19283: gap of 100 bp
* 19284 19976: contig of 693 bp in length
* 19977 20076: gap of 100 bp
* 20077 20781: contig of 705 bp in length
* 20782 20881: gap of 100 bp
* 20882 21590: contig of 709 bp in length
* 21591 21690: gap of 100 bp
* 21691 22393: contig of 703 bp in length
* 22394 22493: gap of 100 bp
* 22494 23204: contig of 711 bp in length
* 23205 23304: gap of 100 bp
* 23305 24016: contig of 712 bp in length
* 24017 24116: gap of 100 bp
* 24117 24824: contig of 708 bp in length
* 24825 24924: gap of 100 bp
* 24925 25634: contig of 710 bp in length
* 25635 25734: gap of 100 bp
* 25735 26414: contig of 680 bp in length
* 26415 26514: gap of 100 bp
* 26515 27215: contig of 701 bp in length
* 27216 27315: gap of 100 bp
* 27316 28025: contig of 710 bp in length
* 28026 28125: gap of 100 bp
* 28126 28835: contig of 710 bp in length
* 28836 28935: gap of 100 bp
* 28936 29643: contig of 708 bp in length
* 29644 29743: gap of 100 bp
* 29744 30451: contig of 708 bp in length
* 30452 30551: gap of 100 bp
* 30552 31252: contig of 701 bp in length
```

```
* 31253 31352: gap of 100 bp
* 31353 32065: contig of 713 bp in length
* 32066 32165: gap of 100 bp
* 32166 32862: contig of 697 bp in length
* 32863 32962: gap of 100 bp
* 32963 33676: contig of 714 bp in length
* 33677 33776: gap of 100 bp
* 33777 34455: contig of 679 bp in length
* 34456 34555: gap of 100 bp
* 34556 35262: contig of 707 bp in length
* 35263 35362: gap of 100 bp
* 35363 36039: contig of 677 bp in length
* 36040 36139: gap of 100 bp
* 36140 36842: contig of 703 bp in length
* 36843 36942: gap of 100 bp
* 36943 37638: contig of 696 bp in length
* 37639 37738: gap of 100 bp
* 37739 38430: contig of 692 bp in length
* 38431 38530: gap of 100 bp
* 38531 39243: contig of 713 bp in length
* 39244 39343: gap of 100 bp
* 39344 40042: contig of 699 bp in length
* 40043 40142: gap of 100 bp
* 40143 40841: contig of 699 bp in length
* 40842 40941: gap of 100 bp
* 40942 41647: contig of 706 bp in length
* 41648 41747: gap of 100 bp
* 41748 42457: contig of 710 bp in length
* 42458 42557: gap of 100 bp
* 42558 43271: contig of 714 bp in length
* 43272 43371: gap of 100 bp
* 43372 44079: contig of 708 bp in length
* 44080 44179: gap of 100 bp
* 44180 44888: contig of 709 bp in length
* 44889 44988: gap of 100 bp
* 44989 45691: contig of 703 bp in length
* 45692 45791: gap of 100 bp
* 45792 46509: contig of 718 bp in length
* 46510 46609: gap of 100 bp
* 46610 47320: contig of 711 bp in length
* 47321 47420: gap of 100 bp
* 47421 48112: contig of 692 bp in length
* 48113 48212: gap of 100 bp
* 48213 48842: contig of 630 bp in length
* 48843 48942: gap of 100 bp
* 48943 49623: contig of 681 bp in length
* 49624 49723: gap of 100 bp
* 49724 50414: contig of 691 bp in length
* 50415 50514: gap of 100 bp
* 50515 51234: contig of 720 bp in length
* 51235 51334: gap of 100 bp
* 51335 52047: contig of 713 bp in length
* 52048 52147: gap of 100 bp
* 52148 52852: contig of 705 bp in length
* 52853 52952: gap of 100 bp
* 52953 53642: contig of 690 bp in length
* 53643 53742: gap of 100 bp
* 53743 54441: contig of 699 bp in length.
```

FEATURES

source

1..54441

/organism="Mus musculus"

Query Match 84.3%; Score 19.4; DB 2; Length 54441;
Best Local Similarity 95.2%; Pred. No. 89;
Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 AACTGGTGGCCATTGACTTC 23

Db 31076 AAATGGTGGCCATTGACTTC 31096

RESULT 12
AC101033/c

Mon Apr 21 10:30:06 2003

AC101033 54441 bp DNA linear HTG 23-NOV-2001
Mus musculus clone RP23-8514, LOW-PASS SEQUENCE SAMPLING.
AC101033
AC101033.1 GI:17059807
HTG: HTGS PHASE0.
Mus musculus.
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
Birren, B., Linton, L., Nussbaum, C. and Lander, E.
1 (bases 1 to 54441)
Mus musculus, clone RP23-8514
Unpublished
2 (bases 1 to 54441)
Birren, B., Linton, L., Nussbaum, C., Lander, E., Ali, A., Allen, N.,
Anderson, S., Barna, N., Bastien, V., Boguslavsky, L., Boukhgalter, B.,
Brown, A., Camarata, J., Campopiano, A., Chang, J., Chazaro, B.,
Choepe, Y., Collangelo, M., Collins, S., Collymore, A., Cook, A.,
Cooke, P., DeArelano, K., Dewar, K., Diaz, J.S., Dodge, S., Faro, S.,
Ferreira, P., Fitzhugh, W., Gage, D., Galagan, J., Gardyna, S.,
Ginde, S., Gord, S., Goyette, M., Graham, L., Grand-pierre, N.,
Hagos, B., Heaford, A., Horton, L., Hulme, W., Iliev, I., Johnson, R.,
Jones, C., Kamat, A., Karatas, A., Kells, C., Lakocque, K.,
Lamaze, R., Landers, T., Lehotzky, J., Levine, R., Liu, G.,
MacDonald, P., Major, J., Marquis, N., Matthews, C.,
McCarthy, M., McEwan, P., McKernan, K., McPheters, R., Meldrim, J.,
Meneus, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C.,
Norbou, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D.,
Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V.,
Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P.,
Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schupbach, R.,
Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N.,
Strauss, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J.,
Topham, K., Travers, M., Travis, N., Trigilio, J., Vassiliev, H.,
Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J., Young, G.,
Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.
Direct Submission
Submitted (23-NOV-2001) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L14840
Center clone name: 85_I_4

* NOTE: This record contains 68 individual
* sequencing reads that have not been assembled into
* contigs. Runs of N are used to separate the reads
* and the order in which they appear is completely
* arbitrary. Low-pass sequence sampling is useful for
* identifying clones that may be gene-rich and allows
* overlap relationships among clones to be deduced.
* However, it should not be assumed that this clone
* will be sequenced to completion. In the event that
* the record is updated, the accession number will
* be preserved.
* 1 689: contig of 689 bp in length
* 690 789: gap of 100 bp
* 790 1570: contig of 781 bp in length
* 1571 1670: gap of 100 bp
* 1671 2370: contig of 700 bp in length
* 2371 2470: gap of 100 bp
* 2471 3175: contig of 705 bp in length
* 3176 3275: gap of 100 bp
* 3276 3943: contig of 668 bp in length
* 3944 4043: gap of 100 bp
* 4044 4742: contig of 699 bp in length

4743 4842: gap of 100 bp
4843 5543: contig of 701 bp in length
5544 5643: gap of 100 bp
5644 6360: contig of 717 bp in length
6361 6460: gap of 100 bp
6461 7157: contig of 697 bp in length
7158 7257: gap of 100 bp
7258 7960: contig of 703 bp in length
7961 8060: gap of 100 bp
8061 8960: contig of 700 bp in length
8961 9566: contig of 706 bp in length
9567 9666: gap of 100 bp
9667 10361: contig of 695 bp in length
10362 10461: gap of 100 bp
10462 11153: contig of 692 bp in length
11154 11233: gap of 100 bp
11234 11964: contig of 711 bp in length
11965 12064: gap of 100 bp
12065 12717: contig of 653 bp in length
12718 12817: gap of 100 bp
12818 13534: contig of 717 bp in length
13535 13634: gap of 100 bp
13635 14333: contig of 699 bp in length
14334 14433: gap of 100 bp
14434 15137: contig of 704 bp in length
15138 15237: gap of 100 bp
15238 15949: contig of 712 bp in length
15950 16049: gap of 100 bp
16050 16760: contig of 711 bp in length
16761 16860: gap of 100 bp
16861 17570: contig of 710 bp in length
17571 18380: contig of 710 bp in length
18381 18480: gap of 100 bp
18481 19183: contig of 703 bp in length
19184 19283: gap of 100 bp
19284 19976: contig of 693 bp in length
19977 20076: gap of 100 bp
20077 20781: contig of 705 bp in length
20782 20881: gap of 100 bp
20882 21590: contig of 709 bp in length
21591 21690: gap of 100 bp
21691 22393: contig of 703 bp in length
22394 22493: gap of 100 bp
22494 23204: contig of 711 bp in length
23205 23304: gap of 100 bp
23305 24016: contig of 712 bp in length
24017 24116: gap of 100 bp
24117 24824: contig of 708 bp in length
24825 24924: gap of 100 bp
24925 25634: contig of 710 bp in length
25635 25734: gap of 100 bp
25735 26414: contig of 680 bp in length
26415 26514: gap of 100 bp
26515 27215: contig of 701 bp in length
27216 27315: gap of 100 bp
27316 28025: contig of 710 bp in length
28026 28125: gap of 100 bp
28126 28835: contig of 710 bp in length
28836 28935: gap of 100 bp
28936 29643: contig of 708 bp in length
29644 29743: gap of 100 bp
29744 30451: contig of 708 bp in length
30452 30551: gap of 100 bp
30552 31252: contig of 701 bp in length
31253 31352: gap of 100 bp
31353 32065: contig of 713 bp in length
32066 32165: gap of 100 bp
32166 32862: contig of 697 bp in length
32863 32962: gap of 100 bp
32963 33676: contig of 714 bp in length
33677 33776: gap of 100 bp

```

* 33777 34455: contig of 679 bp in length
* 34456 34555: gap of 100 bp
* 34556 35262: contig of 707 bp in length
* 35263 35962: gap of 100 bp
* 35963 36039: contig of 677 bp in length
* 36040 36139: gap of 100 bp
* 36140 36842: contig of 703 bp in length
* 36843 36942: gap of 100 bp
* 36943 37638: contig of 696 bp in length
* 37639 37738: gap of 100 bp
* 37739 38430: contig of 692 bp in length
* 38431 38530: gap of 100 bp
* 38531 39243: contig of 713 bp in length
* 39244 39343: gap of 100 bp
* 39344 40042: contig of 699 bp in length
* 40043 40142: gap of 100 bp
* 40143 40841: contig of 699 bp in length
* 40842 40941: gap of 100 bp
* 40942 41647: contig of 706 bp in length
* 41648 41747: gap of 100 bp
* 41748 42457: contig of 710 bp in length
* 42458 42557: gap of 100 bp
* 42558 43271: contig of 714 bp in length
* 43272 43371: gap of 100 bp
* 43372 44079: contig of 708 bp in length
* 44080 44179: gap of 100 bp
* 44180 44888: contig of 709 bp in length
* 44889 44988: gap of 100 bp
* 44989 45691: contig of 703 bp in length
* 45692 45791: gap of 100 bp
* 45792 46509: contig of 718 bp in length
* 46510 46609: gap of 100 bp
* 46610 47320: contig of 711 bp in length
* 47321 47420: gap of 100 bp
* 47421 48112: contig of 692 bp in length
* 48113 48212: gap of 100 bp
* 48213 48842: contig of 630 bp in length
* 48843 48942: gap of 100 bp
* 48943 49623: contig of 681 bp in length
* 49624 49723: gap of 100 bp
* 49724 50414: contig of 691 bp in length
* 50415 50514: gap of 100 bp
* 50515 51234: contig of 720 bp in length
* 51235 51334: gap of 100 bp
* 51335 52047: contig of 713 bp in length
* 52048 52147: gap of 100 bp
* 52148 52852: contig of 705 bp in length
* 52853 52952: gap of 100 bp
* 52953 53642: contig of 690 bp in length
* 53643 53742: gap of 100 bp
* 53743 54441: contig of 699 bp in length.
FEATURES
    source
        1..54441
            /organism="Mus musculus"

```

```

Query Match      84.3%; Score 19.4; DB 2; Length 54441;
Best Local Similarity 95.2%; Pred. No. 89;
Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

QY 3 AACTGTGGCCATTGACCTTC 23
DB 38305 AANTGTGGCCATTGACCTTC 38285

```

```

RESULT 13
AC102692
LOCUS      AC102692
DEFINITION Mus musculus clone RP24-491K18, WORKING DRAFT SEQUENCE, 14
            AC102692
ACCESSION  AC102692
VERSION    AC102692.2 GI:22381681
KEYWORDS   HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE     house mouse.

```

ORGANISM

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE

AUTHORS

1 (bases 1 to 191918)

Birren, B., Nusbaum, C. and Lander, E.

TITLE

Unpublished

REFERENCE

AUTHORS

2 (bases 1 to 191918)

Birren, B., Linton, L., Nusbaum, C., Lander, E., Ali, A., Allen, N., Anderson, S., Barna, N., Bastien, V., Boguslavskiy, L., Boukhgalter, B., Brown, A., Camarata, J., Campolano, A., Chang, J., Chazaro, B., Choepel, Y., Colangelo, M., Collins, S., Collymore, A., Cook, A., Cooke, P., DeArellano, K., Dewar, K., Diaz, J. S., Dodge, S., Faro, S., Ferreira, P., FitzHugh, W., Gage, D., Galagan, J., Gardyna, S., Ginde, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N., Hagos, B., Heaford, A., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kellis, C., LaRocque, K., Lamazares, R., Landers, T., Lehoczy, J., Levine, R., Liu, G., Maclean, C., Macdonald, P., Major, J., Marquis, N., Matthews, C., McCarthy, M., McEwan, P., McKernan, K., McPheeters, R., Meldrim, J., Meneus, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Norbu, C., Norman, C. H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V., Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P., Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schupback, R., Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N., Straus, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M., Travis, N., Trigilio, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

Direct Submission

TITLE

Submitted (23-NOV-2001) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA

REFERENCE

AUTHORS

3 (bases 1 to 191918)

Birren, B., Nusbaum, C., Lander, E., Ali, A., Allen, N., Anderson, S., Barna, N., Bastien, V., Bloom, T., Boguslavskiy, L., Boukhgalter, B., Camarata, J., Chang, J., Chazaro, B., Choepel, Y., Collymore, A., Cook, A., Cooke, P., DeArellano, K., Dewar, K., Diaz, J. S., Dodge, S., Faro, S., Ferreira, P., FitzGerald, M., Gage, D., Galagan, J., Gardyna, S., Gord, S., Graham, L., Grand-Pierre, N., Hagos, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kellis, C., Landers, T., Levine, R., Lindblad-Toh, K., Liu, G., MacLean, C., Macdonald, P., Major, J., Matthews, C., McCarthy, M., Meldrim, J., Meneus, L., Mihova, T., Mienga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., Norman, C. H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Raymond, C., Retta, R., Rise, C., Rogov, P., Roman, J., Roy, A., Schauer, S., Schupback, R., Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Tesfaye, S., Theodore, J., Topham, K., Travers, M., Travis, N., Talamas, J., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

Direct Submission

TITLE

Submitted (21-AUG-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA

REFERENCE

On Aug 21, 2002 this sequence version replaced gi:17060802.

COMMENT

All repeats were identified using RepeatMasker:

Smith, A.P.A. & Green, P. (1996-1997)

http://ftp.genome.washington.edu/RM/RepeatMasker.html

----- Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIBR

Web site: http://www-seq.wi.mit.edu

Contact: sequence_submissions@genome.wi.mit.edu

----- Project Information

Center project name: L19272

Center clone name: 491_K18

----- Summary Statistics

Sequencing vector: Plasmid, n/a; 100% of reads

Chemistry: Dye-terminator Big Dye; 100% of reads

Assembly program: Phrap; version 0.960731

Consensus quality: 187629 bases at least Q40

Consensus quality: 189823 bases at least Q30

```
misc_feature      12383.. assembly fragment"  
/note="assembly fragment"  
BASE COUNT      58327 a 40328 c 39648 g 52315 t    1300 others  
ORIGIN  
  
Query Match      84.3%; Score 19.4; DB 2; Length 191918;  
Best Local Similarity 95.2%; Pred. No. 99;  
Matches 20; Conservative 0; Mismatches 1; Indels 0;  
  
QY              3 AACTGGTGCCATTGAGCTTC 23  
                |||||  
bb 111442 AAATGGTGCCTTCAGCTTC 111462
```

RESULT 14
 AC125105/c
 LOCUS
 DEFINITION
 AC125105
 Mus musculus chromosome UNK clone RP24-220B7, WORKING DRAFT
 SEQUENCE, 60 unordered pieces.
 AC125105.1 GI-21490578
 AC125105.1
 HTG; HTGS_PHASE1; HTGS_DRAFT.
 house mouse.
 SOURCE
 house mouse.
 ORGANISM
 Mus musculus
 Eukaryota; Metazoa; Chordata; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 277892)
 McPherson,J.D. and Waterston,R.H.
 The sequence of Mus musculus clone
 Unpublished
 2 (bases 1 to 277892)
 McPherson,J.D. and Waterston,R.H.
 Direct Submission
 Submitted (20-JUN-2002) Genome Sequencing Center, 4444 Forest Park
 St., Louis, MO 63108, USA

```

----- Genome Center -----
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Center web site: http://genome.wustl.edu/gsc/index.shtml
Contact: submissions@watson.wustl.edu
Project Information
-----
Center project name: M_BB0220B07
----- Summary Statistics -----
Sequencing vector: M13: 0%
Sequencing vector: plasmid: 100%
Chemistry: Dye-primer ET: 0% of reads
Chemistry: Dye-terminator Big Dye: 100% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 248249 bases at least Q40
Consensus quality: 257105 bases at least Q30
Consensus quality: 264370 bases at least Q20
Insert size: 170000; agarose-fp
Insert size: 272299; sum-of-contigs
Quality coverage: 0.00 in Q20 bases; agarose-fp
Quality coverage: 3.87 in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently
* consists of 60 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1 1214: contig of 1214 bp in length
* 1215 1314: gap of unknown length
* 1315 2486: contig of 1172 bp in length
* 2487 2585: gap of unknown length
* 2587 3833: contig of 1253 bp in length
* 3840 3933: gap of unknown length
* 3940 5171: contig of 1232 bp in length

```

Location/Qualifiers

[illegible][illegible]

```

* 5172 5271: gap of unknown length
* 5272 6549: contig of 1278 bp in length
* 6550 6649: gap of unknown length
* 6650 7662: contig of 1013 bp in length
* 7663 9073: gap of unknown length
* 9074 9173: gap of unknown length
* 9174 10263: contig of 1090 bp in length
* 10264 10363: gap of unknown length
* 10364 11990: contig of 1627 bp in length
* 11991 12090: gap of unknown length
* 12091 13171: contig of 1081 bp in length
* 13172 13771: gap of unknown length
* 13772 14786: contig of 1515 bp in length
* 14787 16229: gap of unknown length
* 16230 16330: gap of unknown length
* 16331 18004: contig of 1675 bp in length
* 18005 18105: gap of unknown length
* 18106 19456: contig of 1352 bp in length
* 19457 21458: gap of unknown length
* 21459 21558: gap of unknown length
* 21559 22675: contig of 1117 bp in length
* 22676 22775: gap of unknown length
* 22776 24272: contig of 1497 bp in length
* 24273 24372: gap of unknown length
* 24373 26082: contig of 1710 bp in length
* 26083 26182: gap of unknown length
* 26183 27940: contig of 1758 bp in length
* 27941 28040: gap of unknown length
* 28041 29588: contig of 1548 bp in length
* 29589 29688: gap of unknown length
* 29689 31233: contig of 1545 bp in length
* 31234 31333: gap of unknown length
* 31334 33232: contig of 1898 bp in length
* 33233 33332: gap of unknown length
* 33333 35108: contig of 1777 bp in length
* 35109 35208: gap of unknown length
* 35209 36337: contig of 1428 bp in length
* 36338 36737: gap of unknown length
* 36738 38554: contig of 1818 bp in length
* 38555 40244: contig of 1590 bp in length
* 40245 40344: gap of unknown length
* 40345 42535: contig of 2151 bp in length
* 42536 42635: gap of unknown length
* 42636 44153: contig of 1518 bp in length
* 44154 44253: gap of unknown length
* 44254 45967: contig of 1714 bp in length
* 45968 46067: gap of unknown length
* 46068 47834: contig of 1766 bp in length
* 47835 49139: gap of unknown length
* 49140 49439: gap of unknown length
* 49440 52455: contig of 3016 bp in length
* 52456 52555: gap of unknown length
* 52556 55093: contig of 2538 bp in length
* 55094 55193: gap of unknown length
* 55194 57493: contig of 2300 bp in length
* 57494 57593: gap of unknown length
* 57594 59133: contig of 1540 bp in length
* 59134 60952: contig of 1719 bp in length
* 60953 61053: gap of unknown length
* 61054 63195: contig of 2143 bp in length
* 63196 63295: gap of unknown length
* 63296 65553: contig of 2258 bp in length
* 65554 67988: gap of unknown length
* 67989 68088: gap of unknown length
* 68089 71061: contig of 2973 bp in length
* 71062 71161: gap of unknown length

71162 74202: contig of 3041 bp in length
74203 74302: gap of unknown length
74303 77181: contig of 2879 bp in length
77182 77281: gap of unknown length
77282 80521: contig of 3240 bp in length
80522 80621: gap of unknown length
80622 83147: contig of 2526 bp in length
83148 83247: gap of unknown length
83248 85358: contig of 2111 bp in length
85359 85458: gap of unknown length
85459 89161: contig of 3703 bp in length
89162 89262: gap of unknown length
89263 94379: contig of 5118 bp in length
94380 94479: gap of unknown length
94480 97915: contig of 3436 bp in length
97916 98015: gap of unknown length
98016 103342: contig of 5327 bp in length
103343 103442: gap of unknown length
103443 106304: contig of 2862 bp in length
106305 106404: gap of unknown length
106405 111882: contig of 5478 bp in length
111883 111982: gap of unknown length
111983 116456: contig of 4474 bp in length
116457 116556: gap of unknown length
116557 122640: contig of 6084 bp in length
122641 122741: gap of unknown length
122742 133292: contig of 10552 bp in length
133293 133392: gap of unknown length
133393 144545: contig of 11153 bp in length
144546 144645: gap of unknown length
144646 160567: contig of 15921 bp in length
160568 160666: gap of unknown length
160667 181991: contig of 21325 bp in length
181992 182091: gap of unknown length
182092 216010: contig of 33919 bp in length
216011 216110: gap of unknown length
216111 276917: contig of 60807 bp in length
276918 277018: gap of unknown length
277019 277892: contig of 875 bp in length.

FEATURES
  source
    Location/Qualifiers
      1..277892
        /organism="Mus musculus"
        /db_xref="taxon:10090"
        /chromosome="UNK"
        /clone="RP24-220B7"
      1..1214
        /note="assembly_name:Contig14"
      1315..2486
        /note="assembly_name:Contig18"
      2587..3839
        /note="assembly_name:Contig21"
      3940..5171
        /note="assembly_name:Contig28"
      5272..6549
        /note="assembly_name:Contig33"
      6650..7662
        /note="assembly_name:Contig35"
      7763..9073
        /note="assembly_name:Contig40"
      9174..10263
        /note="assembly_name:Contig41"
      10364..11990
        /note="assembly_name:Contig42"
      12091..13171
        /note="assembly_name:Contig45"
      13272..14786
        /note="assembly_name:Contig46"
      14887..16229
        /note="assembly_name:Contig47"
      16330..18004

Query Match      84.3%  Score 19.4;  DB 2;  Length 277892;
Best Local Similarity 95.4%  Pred. No. 1e+02;

```

Mon Apr 21 10:30:06 2003

Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Search completed: April 18, 2003, 06:25:35
Job time : 582.273 secs

OY 3 AACTGGTGGCCATTGAGCTTC 23
Db 132148 AAATGGTGGCCATTGAGCTTC 132128

RESULT 15
BC013150/c
LOCUS BC013150 1520 bp mRNA linear PRI 29-AUG-2001
DEFINITION Homo sapiens, Similar to hypothetical protein FLJ22595, clone
MGC:17429 IMAGE:4340470, mRNA, complete cds.

ACCESSION BC013150
VERSION BC013150.1 GI:15341924
KEYWORDS MGC.

SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1520)

REFERENCE Strausberg, R.
AUTHORS Direct Submission
TITLE Submitted (27-AUG-2001) National Institutes of Health, Mammalian
JOURNAL Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA

REMARK NIH-MGC Project URL: <http://mgc.nci.nih.gov>
COMMENT Contact: MGC help desk
Email: cgaps-remail.nih.gov
Tissue Procurement: Lou Staudt
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Baylor College of Medicine Human Genome
Sequencing Center
Center code: BCM-HGSC
Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>
Contact: villalon@bcm.tmc.edu.
Villalon, D.K., Luna, R.A., Hale, S.M., Hulyk, S., Lu, X., Garcia,
A.M., Holloway, M., Telford, B., Hodgson, A., Bouck, J., Yu, W.,
Muzny, D.M., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAX Plate: 26 Row: 1 Column: 7.

FEATURES
source
1..1520
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="MGC:17429 IMAGE:4340470"
/tissue_type="lymph, lymphoma"
/clone_lib="NIH_MGC_85"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6"
115..705
/codon_start=1
/product="Similar to hypothetical protein FLJ22595"
/protein_id="AAH13150.1"
/db_xref="GI:15341925"
/translation="MGSVNSGHKAEQVVMGLDSAGKTTLLYKLGKQHVETLPTV
GFNVEPLKAPGVSLTWDVGQAPLRASWDYLEGTDILVYVLDSTDEARLPESAAE
LTVELNDPNMAGVPEFLVLAQKQAPDPLPKIRNRLSLERFQDHCHELRCGSALTGE
GLPEALQSLNLSLQKRSQCLQARAHGAERGDGSKRS"
BASE COUNT 414 a 372 c 431 g 303 t

ORIGIN
Query Match 81.7%; Score 18.8; DB 9; Length 1520;
Best Local Similarity 90.9%; Pred. No. 1.3e+02;
Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
OY 1 CCAACTGGTGGCCATTGAGCTT 22
Db 226 CCAGCTGGTGGCCCTTCAGCTT 205

GenCore version 5.1.4.p5 4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 18, 2003, 04:12:35 ; Search time 102.455 Seconds
(without alignments)
505.551 Million cell updates/sec

Title: US-09-270-437D-12

Perfect score: 23
Sequence: 1 ccaactgggtgcccattcagcttc 23

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_101002:*

- 1: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:*
- 2: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:*
- 3: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:*
- 4: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:*
- 5: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:*
- 6: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:*
- 7: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:*
- 8: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:*
- 9: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:*
- 10: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:*
- 11: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:*
- 12: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:*
- 13: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:*
- 14: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:*
- 15: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:*
- 16: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:*
- 17: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:*
- 18: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:*
- 19: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:*
- 20: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:*
- 21: /SID22/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:*
- 22: /SID22/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:*
- 23: /SID22/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:*
- 24: /SID22/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	23	100.0	23	AAZ36158	PCR primer for DNA
2	23	100.0	1708	AAZ36151	DNA encoding cance
3	23	100.0	1946	21 AAZ36153	An alternative for
4	18.8	81.7	415	22 AAK58003	Human immune/haema
5	18.8	81.7	1109	21 AAA49190	CDNA encoding huma
6	18.8	81.7	2568	23 ABL28587	Drosophila melanog
7	18.8	81.7	2834	23 ABL21504	Drosophila melanog
8	18.8	81.7	3722	22 AAK71737	Human immune/haema
9	18.8	81.7	4870	23 ABL28586	Drosophila melanog

C	10	17.8	77.4	3628	24	ABK63547	Rat sequence diffc
	11	17.8	77.4	8212	24	ABI99884	Mouse ischaemic co
	12	17.2	74.8	306	13	AAQ24805	Region 3' of human
	13	17.2	74.8	1019	18	AAV75063	Staphylococcus aur
	14	17.2	74.8	1203	13	AAQ24809	Mutant human D3 do
	15	17.2	74.8	1261	21	AAQ9614	Human D2 receptor
	16	17.2	74.8	1272	13	AAQ4803	Human D3 dopaminer
	17	17.2	74.8	1558	13	AAQ27988	Protease from S. A
	18	17.2	74.8	2325	16	AAQ04156	Enzyme Q36 coding
	19	17.2	74.8	3073	16	AAQ04158	Enzyme Q36 coding
C	20	17.2	74.8	4172	23	ABL12003	Drosophila melanog
	21	17.2	74.8	6781	23	ABL12002	Drosophila melanog
	22	17.2	74.8	17170	22	AAK77880	Human immune/haema
	23	17.2	74.8	17173	22	AAK77881	Human immune/haema
C	24	17.2	74.8	39159	23	ABL30024	Drosophila melanog
C	25	16.8	73.0	258	24	ABQ20709	Human genome-deriv
	26	16.8	73.0	309	23	ABQ61599	Human genome-deriv
C	27	16.8	73.0	454	24	ABQ8292	Human prostate exp
	28	16.8	73.0	7506	23	ABV24912	Human prostate exp
C	29	16.6	72.2	510	22	AAK11576	Human brain expres
	30	16.6	72.2	708	24	ABN67249	Streptococcus poly
C	31	16.6	72.2	746	22	AAH99822	Human protein enco
C	32	16.6	72.2	1645	23	ABL09057	Drosophila melanog
C	33	16.6	72.2	1987	23	ABL20943	Drosophila melanog
C	34	16.6	72.2	2224	20	AAZ10617	CDNA encoding a mu
C	35	16.6	72.2	2305	23	ABV24325	Human prostate exp
C	36	16.6	72.2	2305	23	ABV24337	Human prostate exp
C	37	16.6	72.2	2305	23	ABV24365	Human prostate exp
C	38	16.6	72.2	2379	23	ABL16853	Drosophila melanog
C	39	16.6	72.2	2496	24	ABQ67860	Listeria innocua D
C	40	16.6	72.2	2496	24	ABQ69432	Listeria innocua D
C	41	16.6	72.2	2575	24	ABQ70592	Listeria monocytog
C	42	16.6	72.2	2830	23	ABL17257	Drosophila melanog
C	43	16.6	72.2	2971	23	ABL13725	Drosophila melanog
	44	16.6	72.2	2980	23	ABL17175	Drosophila melanog
	45	16.6	72.2	4212	23	ABL12981	Drosophila melanog

ALIGNMENTS

RESULT 1
AAZ36158
ID AAZ36158 standard; DNA; 23 Bp.
XX AC AAZ36158;
XX AC AAZ36158;
XX DT 11-FEB-2000 (first entry)
DE DE PCR primer for DNA encoding cancer associated antigen KOC-2.
XX XX Cancer associated antigen; KOC-1; cancer; vaccine; CT7; PCR primer; 89.
XX OS Synthetic.
OS OS Homo sapiens.
XX XX WO954738-A1.
XX XX 28-OCT-1999.
XX PF 16-MAR-1999; 99WO-US05766.
XX PR 17-APR-1998; 98US-0061709.
XX XX (LUDW-) LUDWIG INST CANCER RES.
XX PA Chen Y, Gure A, Tsang S, Stockert E, Jager E, Knuth A, Old L;
XX PI WPI; 2000-013284/01.
XX DR Nucleotides representing cancer-associated genes, used to develop
XX PT products for the diagnosis, monitoring and treatment of cancers -

RESULT 4
AAKS8003/c
ID AAKS8003 standard; CDNA; 415 BP.
AC AAKS8003;
DT 06-NOV-2001 (first entry)
XX Human immune/haematopoietic antigen encoding cDNA SEQ ID NO:3063.
DE Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
KW Cytostatic; gene therapy; vaccine; metastasis; ss.
XX Homo sapiens.
XX WO200157182-A2.
FN 09-AUG-2001.
PD 17-JAN-2001; 2001WO-US01354.
PF 31-JAN-2000; 2000US-0179065.
XX 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226568.
PR 22-AUG-2000; 2000US-0226688.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 21-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 25-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 26-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 13-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0244674.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.

us-09-270-437d-12.rng

Mon Apr 21 10:30:08 2003

```

PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
XX (HUMA-) HUMAN GENOME SCI INC.
XX Rosen CA, Barash SC, Ruben SM;
XX P-PSDB; AAM85222.
XX WPI; 2001-483426/52.
XX P-PSDB; AAM85222.
XX Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
XX useful for preventing, diagnosing and/or treating cancers and
XX metastasis -
XX Claim 1; SEQ ID NO 3063; 3071pp + Sequence Listing; English.
XX AAK54951 to AAK64702 encode the human immune/hematopoietic antigen (I)
XX amino acid sequences given in AAM82170 to AAM91921. (I) have cytostatic
XX activity, and can be used in gene therapy and vaccine production. (I)
XX proteins and polynucleotides may be used in the prevention, diagnosis and
XX treatment of diseases associated with inappropriate (I) expression. For
XX example, they may be used to treat disorders associated with decreased
XX expression by rectifying mutations or deletions in a patient's genome
XX that affect the activity of (I) by expressing inactive proteins or to
XX supplement the patient's own production of (I). Additionally, (I)
XX polynucleotides may be used to produce the secreted (I), by inserting
XX the nucleic acids into a host cell and culturing the cell to express the
XX protein. (I) proteins and polynucleotides may be used to prevent,
XX diagnose and treat immune/hematopoietic-related diseases, especially
XX cancers and cancer metastases of hematopoietic-derived cells. AAK64703
XX to AAK87694 represent human immune/hematopoietic antigen genomic
XX sequences from the present invention. AAK54942 to AAK54950 and AAM82169
XX represent sequences used in the exemplification of the present invention.
XX
XX Sequence 415 BP; 92 A; 114 C; 130 G; 76 T; 3 other;
XX
Query Match 81.7%; Score 18.8; DB 22; Length 415;
Best Local Similarity 90.9%; Pred. No. 28;
Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Oy 1 CCAACTGGTGGCCATTTCAGCTT 22
Db 226 CCAGCTGGTGGCCCTTCAGCTT 205
RESULT 5
AA49190/c
ID AAA49190 standard; cDNA; 1109 BP.
XX AAA49190;
XX
XX 03-NOV-2000 (first entry)
XX
XX cDNA encoding human GTPase associated protein-20.
XX
XX Guanine nucleotide binding protein; GTP-binding protein; G-protein;
XX GTPase; GTPase associated protein; GTPAP; cell proliferation;
XX autoimmune; inflammatory; immune system disorder; cancer; AIDS;
XX acquired immune deficiency syndrome; asthma; atherosclerosis;
XX arthritis; systemic lupus erythematosus; psoriasis; human; ss.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX

```

```

CDS
FT 127..717
FT /*tag= a
FT /product= GTPAP20
XX
XX WO2000031263-A2.
XX
XX 02-JUN-2000.
XX
XX 23-NOV-1999; 99WO-US28013.
XX
XX 23-NOV-1998; 98US-0109592.
XX
XX 04-FEB-1999; 99US-0118610.
XX
XX 06-APR-1999; 99US-0127990.
XX
XX (INCY-) INCYTE PHARM INC.
XX
XX Hillman JL, Tang YT, Bandman O, Lal P, Yue H, Lu DAM, Baughn MR;
XX Yang J, Azimzai Y;
XX
XX WPI; 2000-400073/34.
XX
XX P-PSDB; AAY99669.
XX
XX Human GTPase associated proteins, polynucleotides, and antibodies,
XX useful for diagnosing, preventing and treating various diseases such as
XX atherosclerosis, cancer, acquired immune deficiency syndrome (AIDS),
XX asthma, and autoimmune diseases -
XX
XX Claim 9; Page 136; 144pp; English.
XX
XX Human cDNA libraries from various tissues were screened for GTPase
XX associated proteins (GTPAP). The present sequence is cDNA encoding
XX human GTPAP-20. This sequence was derived from a cDNA library of the
XX wrist synovial membrane tissue from a female with rheumatoid arthritis.
XX This protein is expressed in reproductive, nervous and
XX gastrointestinal tissue. The GTPAP proteins may be used to define
XX agonists and antagonists of GTPAP activity and to generate antibodies
XX to GTPAP. This means the GTPAP proteins may be useful for treatment or
XX prevention of diseases associated with GTPAP such as cell proliferation
XX disorders, autoimmune disorders, inflammatory disorders, immune system
XX disorders, cancer, AIDS, asthma, atherosclerosis, arthritis, systemic
XX lupus erythematosus and psoriasis.
XX
XX Sequence 1109 BP; 264 A; 297 C; 334 G; 214 T; 0 other;
XX
Query Match 81.7%; Score 18.8; DB 21; Length 1109;
Best Local Similarity 90.9%; Pred. No. 32;
Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Oy 1 CCAACTGGTGGCCATTTCAGCTT 22
Db 238 CCAGCTGGTGGCCCTTCAGCTT 217
RESULT 6
ABL28587/c
ID ABL28587 standard; DNA; 2568 BP.
XX ABL28587;
XX
XX 26-MAR-2002 (first entry)
XX
XX Drosophila melanogaster genomic polynucleotide SEQ ID NO 37234.
XX
XX Drosophila; developmental biology; cell signalling; insecticide;
XX pharmaceutical; gene; ds.
XX
XX Drosophila melanogaster.
XX
XX WO200171042-A2.
XX
XX 27-SEP-2001.
XX
XX 23-MAR-2001; 2001WO-US09231.
XX

```

XX PR 23-MAR-2000; 2000US-191637P.
PR PR 11-JUL-2000; 2000US-0614150.
XX PA (PEKE) PE CORP NY.
XX PI Venter JC, Adams M, Li PWD, Myers EW;
XX WPI; 2001-656860/75.
DR New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions -
XX Claim 1; SEQ ID NO 37234; 21pp + Sequence Listing; English.
XX The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins
CC (ABB57737-ABB72072).
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX SQ Sequence 2568 BP; 686 A; 575 C; 664 G; 643 T; 0 other;
XX Query Match 81.7%; Score 18.8; DB 23; Length 2568;
Best Local Similarity 90.9%; Pred. No. 37;
Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 2 CAACTGGTGCCATTGAGCTTC 23
DB 1965 CAACTGGTGCCACTGAGCTTC 1944
RESULT 7
ABL21504
ID ABL21504 standard; DNA; 2834 BP.
XX AC ABL21504;
XX DT 26-MAR-2002 (first entry)
XX DE Drosophila melanogaster genomic polynucleotide SEQ ID NO 15985.
XX KW Drosophila; developmental biology; cell signalling; insecticide;
XX KW pharmaceutical; gene; ds.
XX OS Drosophila melanogaster.
XX PN WO200171042-A2.
XX PD 27-SEP-2001.
XX PF 23-MAR-2001; 2001WO-US09231.
XX PR 23-MAR-2000; 2000US-191637P.
XX PR 11-JUL-2000; 2000US-0614150.
XX PA (PEKE) PE CORP NY.
XX PI Venter JC, Adams M, Li PWD, Myers EW;
XX WPI; 2001-656860/75.
XX New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions -
XX Claim 1; SEQ ID NO 37234; 21pp + Sequence Listing; English.
XX The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins
CC (ABB57737-ABB72072).
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX SQ Sequence 2568 BP; 686 A; 575 C; 664 G; 643 T; 0 other;
XX Query Match 81.7%; Score 18.8; DB 23; Length 2568;
Best Local Similarity 90.9%; Pred. No. 37;
Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 2 CAACTGGTGCCATTGAGCTTC 23
DB 1965 CAACTGGTGCCACTGAGCTTC 1944

PS Claim 1; SEQ ID NO 15985; 21pp + Sequence Listing; English.
XX The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins
CC (ABB57737-ABB72072).
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX SQ Sequence 2834 BP; 747 A; 689 C; 625 G; 773 T; 0 other;
XX Query Match 81.7%; Score 18.8; DB 23; Length 2834;
Best Local Similarity 90.9%; Pred. No. 37;
Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 2 CAACTGGTGCCATTGAGCTTC 23
DB 2773 CAACTGGTGCCACTGAGCTTC 2794
RESULT 8
AAK71737/c
ID AAK71737 standard; DNA; 3722 BP.
XX AC AAK71737;
XX DT 06-NOV-2001 (first entry)
XX DE Human immune/haematopoietic antigen genomic sequence SEQ ID NO:26549.
XX KW Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
XX KW cytostatic; gene therapy; vaccine; metastasis; ds.
XX OS Homo sapiens.
XX PN WO200157182-A2.
XX PD 09-AUG-2001.
XX PF 17-JAN-2001; 2001WO-US01354.
XX PR 31-JAN-2000; 2000US-0179065.
XX PR 04-FEB-2000; 2000US-0180628.
XX PR 24-FEB-2000; 2000US-0184664.
XX PR 02-MAR-2000; 2000US-0186350.
XX PR 16-MAR-2000; 2000US-0189874.
XX PR 17-MAR-2000; 2000US-0190076.
XX PR 18-APR-2000; 2000US-0198123.
XX PR 19-MAY-2000; 2000US-0205515.
XX PR 07-JUN-2000; 2000US-0209467.
XX PR 28-JUN-2000; 2000US-0214886.
XX PR 30-JUN-2000; 2000US-0215135.
XX PR 07-JUL-2000; 2000US-0216647.
XX PR 07-JUL-2000; 2000US-0216880.
XX PR 11-JUL-2000; 2000US-0217487.
XX PR 11-JUL-2000; 2000US-0217496.
XX PR 14-JUL-2000; 2000US-0218290.
XX PR 26-JUL-2000; 2000US-0220963.
XX PR 26-JUL-2000; 2000US-0220964.
XX PR 14-AUG-2000; 2000US-0224518.
XX PR 14-AUG-2000; 2000US-0224519.
XX PR 14-AUG-2000; 2000US-0225213.
XX PR 14-AUG-2000; 2000US-0225214.
XX PR 14-AUG-2000; 2000US-0225266.
XX PR 14-AUG-2000; 2000US-0225267.
XX PR 14-AUG-2000; 2000US-0225268.
XX PR 14-AUG-2000; 2000US-0225270.
XX PR 14-AUG-2000; 2000US-0225447.

PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226868.
PR 22-AUG-2000; 2000US-0227182.
PR 22-AUG-2000; 2000US-0227009.
PR 23-AUG-2000; 2000US-0228924.
PR 30-AUG-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 01-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 08-SEP-2000; 2000US-0231968.
PR 12-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241121.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246529.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246533.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249219.
PR 17-NOV-2000; 2000US-0249220.
PR 17-NOV-2000; 2000US-0249221.
PR 17-NOV-2000; 2000US-0249222.
PR 17-NOV-2000; 2000US-0249223.
PR 17-NOV-2000; 2000US-0249224.
PR 17-NOV-2000; 2000US-0249225.
PR 17-NOV-2000; 2000US-0249226.
PR 17-NOV-2000; 2000US-0249227.
PR 17-NOV-2000; 2000US-0249228.
PR 17-NOV-2000; 2000US-0249229.
PR 17-NOV-2000; 2000US-0249230.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 06-DEC-2000; 2000US-0251856.
PR 06-DEC-2000; 2000US-0251858.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
XX (HUMA-) HUMAN GENOME SCI INC.
PA Rosen CA, Barash SC, Ruben SM;
XX WPI; 2001-483426/52.
XX Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
PT useful for preventing, diagnosing and/or treating cancers and
PT metastasis -
XX Disclosure; SEQ ID NO 26549; 307lpp + Sequence Listing; English.
XX AAK54951 to AAK64702 encode the human immune/hematopoietic antigen (I)
XX amino acid sequences given in AAK82170 to AAK91921. (I) have cytostatic
XX activity, and can be used in gene therapy and vaccine production. (I)
XX proteins and polynucleotides may be used in the prevention, diagnosis and
XX treatment of diseases associated with inappropriate (I) expression. For
XX example, they may be used to treat disorders associated with decreased
XX expression by rectifying mutations or deletions in a patient's genome
XX that affect the activity of (I) by expressing inactive proteins or to
XX supplement the patients own production of (I). Additionally, (I)
XX polynucleotides may be used to produce the secreted (I), by inserting the
XX the nucleic acids into a host cell and culturing the cell to express the
XX protein. (I) proteins and polynucleotides may be used to prevent,
XX diagnose and treat immune/hematopoietic-related diseases, especially
XX cancers and cancer metastases of haematopoietic-derived cells. AAK64703
XX to AAK87694 represent human immune/hematopoietic antigen genomic
XX sequences from the present invention. AAK54942 to AAK54950 and AAK82169
XX represent sequences used in the exemplification of the present invention.
XX Sequence 3722 BP; 990 A; 887 C; 926 G; 919 T; 0 other;
Query Match 81.7%; Score 18.8; DB 22; Length 3722;
Best Local Similarity 90.9%; Pred. NO. 39;
Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CCAACTGGTGGCCATTCAGCTT 22
||| ||||| |||||
Db 2041 CCAGCTGGTGGCCCTTCAGCTT 2020

RESULT 9

ABL28586
ID ABL28586 standard; DNA; 4870 BP.
XX AC ABL28586;
XX DT 26-MAR-2002 (first entry)

XX DE Drosophila melanogaster genomic polynucleotide SEQ ID NO 37231.
XX KW Drosophila; developmental biology; cell signalling; insecticide;
XX KW pharmaceutical; gene; ds.
XX OS Drosophila melanogaster.

XX PN WO200171042-A2.
XX PD 27-SEP-2001.

XX PF 23-MAR-2001; 2001WO-US09231.
XX PR 23-MAR-2000; 2000US-191637P.
XX PR 11-JUL-2000; 2000US-0614150.

XX PA (PEXE) PE CORP NY.
XX PI Venter JC, Adams M, Li PWD, Myers EW;
XX PI WPI; 2001-656860/75.

XX PT New isolated nucleic acid detection reagent for detecting 1000 or more
XX PT genes from Drosophila and for elucidating cell signalling and cell-cell
XX PT interactions -

XX PS Claim 1; SEQ ID NO 37231; 21pp + Sequence Listing; English.
XX CC The invention relates to an isolated nucleic acid detection reagent
XX CC capable of detecting 1000 or more genes from Drosophila. The invention is
XX CC useful in developmental biology and in elucidating cell signalling and
XX CC cell-cell interactions in higher eukaryotes for the development of
XX CC insecticides, therapeutics and pharmaceutical drugs. The invention
XX CC discloses genomic DNA sequences (AB116176-AB120511), expressed DNA
XX CC sequences (AB101840-AB116175) and the encoded proteins
XX CC (AB157737-AB172072).

XX CC The sequence data for this patent did not form part of the printed
XX CC specification, but was obtained in electronic format directly from WIPO
XX CC at ftp.wipo.int/pub/published_pct_sequences.

XX SQ Sequence 4870 BP; 1239 A; 1201 C; 1083 G; 1347 T; 0 other;
Query Match 81.7%; Score 18.8; DB 23; Length 4870;
Best Local Similarity 90.9%; Pred. No. 41;
Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 CCAACTGGTGGCCATTCAGCTTC 23
||| ||||| |||||
Db 1662 CCACTGGTGGTCCACTCAGCTTC 1683

RESULT 10

ABK63547/c
ID ABK63547 standard; cDNA; 3628 BP.
XX AC ABK63547;
XX DT 18-JUN-2002 (first entry)

XX XX Rat sequence differentially expressed in response to a hepatotoxin #1454.

XX KW Rat; ss; hepatotoxin; expressed sequence tag; EST; drug screening;
XX KW differential expression; centrilobular necrosis; steatosis.
XX OS Rattus norvegicus.
XX PN WO200210453-A2.
XX PD 07-FEB-2002.
XX PF 30-JUL-2001; 2001WO-US23872.
XX PR 31-JUL-2000; 2000US-222040P.
XX PR 02-NOV-2000; 2000US-244880P.
XX PR 11-MAY-2001; 2001US-290029P.
XX PR 15-MAY-2001; 2001US-290645P.
XX PR 22-MAY-2001; 2001US-292336P.
XX PR 06-JUN-2001; 2001US-295798P.
XX PR 13-JUN-2001; 2001US-297457P.
XX PR 19-JUN-2001; 2001US-298884P.
XX PR 09-JUL-2001; 2001US-303459P.
XX PA (GENE-) GENE LOGIC INC.
XX PI Mendrick D, Porter MW, Johnson KR, Castle AL, Elashoff MR;
XX PI WPI; 2002-241625/29.
XX PT Predicting toxic effects of compounds or the progression of these toxic
XX PT effects by determining the changes in gene expression in tissues or
XX PT cells exposed to the toxin and comparing these to gene expression in
XX PT unexposed tissues or cells -
XX PS Claim 1; Seq ID No 1454; 239pp; English.
XX CC The invention relates to methods for predicting toxic effects of
XX CC compounds or the progression of these toxic effects by determining the
XX CC global changes in gene expression in tissues or cells exposed to the
XX CC toxin and comparing these to gene expression in unexposed tissues or
XX CC cells. Also included are methods of predicting at least one toxic
XX CC effect of a compound or progression of a toxic effect, preferably the
XX CC hepatotoxicity of a compound, comprising detecting the level of
XX CC expression in a tissue or cell sample exposed to the compound of two or
XX CC more genes listed in the specification, where differential expression of
XX CC the genes is indicative of at least one toxic effect or progression.
XX CC The method can also be used to identify an agent which modulates the
XX CC toxic response and predict cellular pathways that a compound modulates
XX CC in a cell. The methods utilise a set of at least two probes (on a solid
XX CC support in kit form), where each of the probes comprises a sequence that
XX CC specifically hybridises to a gene listed in the specification, a computer
XX CC system comprising a database containing information identifying the
XX CC expression level in a tissue or cell sample exposed to a hepatotoxin of a
XX CC set of genes comprising at least two genes listed in the specification,
XX CC and a user interface to view the information used to present information,
XX CC identifying the expression level in a tissue or cell of at least one gene
XX CC listed in the specification. The method is useful for elucidating global
XX CC changes in gene expression and for identifying toxicity markers in
XX CC tissues or cell exposed to a known toxin. The genes may be used as
XX CC toxicity markers in drug screening and toxicity assays. The genes and
XX CC gene expression information may be used as diagnostic markers for the
XX CC prediction or identification of the physiological state of tissue or cell
XX CC sample that has been exposed to a compound or agent. Hepatotoxicity
XX CC is characterised by centrilobular necrosis and steatosis. The present
XX CC sequence is an expressed sequence tag (EST) or cDNA derived from a gene
XX CC which is differentially expressed in response to a hepatotoxic agent.
XX SQ Sequence 3628 BP; 1035 A; 698 C; 838 G; 1057 T; 0 other;
Query Match 77.4%; Score 17.8; DB 24; Length 3628;
Best Local Similarity 90.5%; Pred. No. 1.2e+02;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Mon Apr 21 10:30:08 2003

us-09-270-437d-12.rng

```

XX DE Region 3' of human D3 dopaminergic receptor ORF.
XX KW G-protein-coupled receptors; tumour; psychosis; Parkinson's disease;
XX KW arterial hypertension; hypothalamo-hypophyseal axis;
XX KW open reading frame; ss.
XX OS Homo sapiens.
XX FH Key Location/Qualifiers
XX FT misc_feature 47..240
XX FT FT /*tag= a
XX FT FT /note= "corresponds to nucleotides
XX FT FT 1013-1206 of AAQ24803"
XX PN WO9207937-A.
XX PD 14-MAY-1992.
XX PF 16-OCT-1991; 91WO-FR00810.
XX PF 06-NOV-1990; 90FR-0013731.
XX PR (INRM ) INSERM INST NAT SANTE RECH MEDICALE.
XX PA Giros B, Martres MP, Schwartz JC, Sokoloff P;
XX PI WPI; 1992-183673/22.
XX DR Human D3 dopaminergic receptor polypeptide(s) - for in-vitro
XX PT screening of new dopaminergic drugs and diagnosis and treatment
XX PT of various psychiatric, cardiovascular, neurological and
XX PT neuro-endocrinological disorders
XX PS Claim 4; Fig 4; 51pp; French.
XX CC A Sau3A human genomic DNA bank was screened with two rat D3 receptor
XX CC probes to identify open reading frames. The rat D3 probe #2
XX CC containing the coding sequence for the C-terminal part and
XX CC transmembrane domains VI and VII of rat D3 identified 1 positive
XX CC clone. The sequence given here corresponds to part of the 3',
XX CC region of the human D3 gene. This sequence is one of the preferred
XX CC fragments able to code for a polypeptide with the properties of
XX CC the human D3 receptor. The sequence is also useful as a probe to
XX CC diagnose e.g. pathological expression of the D3 receptor,
XX CC polymorphism of the D3 receptor gene or point mutations.
XX CC See AAQ24803-9.
XX SQ Sequence 306 BP; 63 A; 92 C; 69 G; 82 T; 0 other;
XX Query Match 74.8%; Score 17.2; DB 13; Length 306;
XX Best Local Similarity 86.4%; Pred. No. 1.5e+02;
XX Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 CCAACTGGTGGCCATTGAGTTC 22
    |||||
DB 4 CCAATGGTGGCCATTGAGTTC 25
    |||||

RESULT 13
AAV75063
ID AAV75063 standard; DNA; 1019 BP.
XX AC AAV75063;
XX DT 16-MAR-1999 (first entry)
XX DE Staphylococcus aureus contig SEQ ID #752.
XX KW Computer readable medium; vaccine; S.aureus infection; immunodetection;
XX KW cellulitis; eyelid infection; food poisoning; osteomyelitis; therapy;
XX KW skin infection; surgical wound infection; scalded skin syndrome;
XX KW toxic shock syndrome; ds.

```

```

Db 1355 AGCTGGAGGCCATTGAGTTC 1335
    |||||
RESULT 11
ABI99884
ID ABI99884 standard; cDNA; 8212 BP.
XX AC ABI99884;
XX DT 07-MAR-2002 (first entry)
XX DE Mouse ischaemic condition related cDNA sequence SEQ ID NO:1027.
XX KW Mouse; ischaemia; compressive ischaemia; occlusive ischaemia;
XX KW vasospastic ischaemia; ischaemic condition; ischaemic disease; ss.
XX OS Mus musculus.
XX PN WO200188188-A2.
XX PD 22-NOV-2001.
XX PF 18-MAY-2001; 2001WO-JP04192.
XX PF 18-MAY-2000; 2000JP-0145977.
XX PR (UYN1-) UNIV NIHON SCHOOL JURIDICAL PERSON.
XX PA Ishikawa K, Asai S, Takahashi Y, Nagata T, Ishii Y;
XX PI WPI; 2002-034733/04.
XX DR Examining the ischemic condition (e.g. occlusive ischemia) by measuring
XX PT expression levels of particular genes defined in the specification or
XX PT by determining the expression profile of a gene group comprising these
XX PT genes -
XX PS Claim 2; Page 2592-2596; 2690pp; English.
XX CC The present invention describes a method for examining ischemic
XX CC conditions, comprising measuring the expression levels of particular
XX CC genes (1) in a test sample or determining the expression profile of a
XX CC gene group in the sample comprising genes selected from (I). The method
XX CC is useful for examining the ischemic condition (e.g. compressive
XX CC ischaemia, occlusive ischaemia or vasospastic ischaemia) by measuring
XX CC expression levels of particular genes (ABI99202 to ABI99912, encoding
XX CC the protein sequences in ABB57020 to ABB57374) or by determining the
XX CC expression profile of a gene group comprising these genes. The
XX CC expression levels or expression profiles produced by these genes are
XX CC used as an indicator when screening for ischaemic condition-improving
XX CC drugs or therapeutics for ischaemic diseases. ABI99913 and ABI99914
XX CC represent PCR primers for a mouse ischaemic condition related sequence,
XX CC which are used in the exemplification of the present invention.
XX SQ Sequence 8212 BP; 1663 A; 2301 C; 2300 G; 1948 T; 0 other;
XX Query Match 77.4%; Score 17.8; DB 24; Length 8212;
XX Best Local Similarity 90.5%; Pred. No. 1.3e+02;
XX Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 AACTGGTGGCCATTGAGTTC 23
    |||||
DB 1006 AACTGGTGGCCATTGAGTTC 1026
    |||||

RESULT 12
AAQ24805
ID AAQ24805 standard; DNA; 306 BP.
XX AC AAQ24805;
XX DT 19-NOV-1992 (first entry)

```


XX OS Staphylococcus aureus.

XX FH Key Location/Qualifiers

XX FT misc_feature 901..960

XX FT /*tag= a

XX FT /note= "these bases represent a line of missing text in the sequence listing in the specification. They are included to maintain the nucleotide numbering given in the specification for this DNA sequence"

XX FT

XX PN EP786519-A2.

XX XX

XX PD 30-JUL-1997.

XX XX

XX PF 07-JAN-1997; 97EP-0100117.

XX XX

XX PR 05-JAN-1996; 96US-0009861.

XX XX

XX PA (HUMA-) HUMAN GENOME SCI INC..

XX PI Barash SC, Choi GH, Dillon PJ, Fannon MR, Kunsch CA; Rosen CA;

XX PI

XX XX WPI; 1997-374922/35.

XX DR

XX XX Polynucleotide(s) and proteins derived from Staphylococcus aureus - stored on computer readable medium and used in the production of anti-S.aureus vaccines

XX PT

XX PS Claim 1; Page 1646-1647; 327lpp; English.

XX CC This sequence represents one of 5191 Staphylococcus aureus DNA sequences of the invention. The DNA sequences are recorded on a computer readable medium, preferably selected from a floppy or hard disk, random access memory (RAM), read-only memory (ROM) or CD-ROM. Homology searches using the S.aureus DNA sequences allows putative functions to be assigned so that protein-encoding or regulatory regions of commercial, therapeutic or industrial importance can be obtained. Specifically, sequences which are likely to encode antigens have been identified and these polypeptides can be used in a vaccine composition against S.aureus infection. The polypeptides can also be used in a kit for the immunodetection of S.aureus in a sample. S.aureus is implicated in numerous human diseases, including cellulitis, eyelid infections, food poisoning, osteomyelitis, skin and surgical wound infections, scalded skin syndrome, toxic shock syndrome, etc. Organisms transformed with the DNA sequences can be used for recombinant production of the polypeptides. The new DNA sequences (and their fragments) are useful as primers or probes for isolating homologues of any of the S.aureus DNA sequences contained on the computer readable medium.

XX SQ Sequence 1019 BP; 364 A; 194 C; 159 G; 240 T; 62 other;

Query Match 74.8%; Score 17.2; DB 18; Length 1019;

Best Local Similarity 86.4%; Pred. No. 1.8e+02;

Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 CAACCTGGTGGCCATTGACCTTC 23

DB 640 CAACCTGGTGGTAATTCAGGTT 661

RESULT 14

AAQ24809

ID AAQ24809 standard; DNA; 1203 BP.

XX AC AAQ24809;

XX XX

XX DT 19-NOV-1992 (first entry)

XX DE Mutant human D3 dopaminergic receptor.

XX XX

KW G-protein-coupled receptors; tumour; psychosis; Parkinson's disease;

KW arterial hypertension; hypothalamo-hypophyseal axis; ss.

XX OS Homo sapiens.

XX FH Key Location/Qualifiers

XX FT mutation 25

XX FT /*tag= a

XX FT /note= "A is replaced by G"

XX FT mutation 51

XX FT /*tag= b

XX FT /note= "A is replaced by G"

XX FT mutation 236

XX FT /*tag= c

XX FT /note= "C is replaced by T"

XX FT mutation 565

XX FT /*tag= d

XX FT /note= "G is replaced by A"

XX FT mutation 876

XX FT /*tag= e

XX FT /note= "T is replaced by C"

XX FT mutation 1190

XX FT /*tag= f

XX FT /note= "T is replaced by C"

XX PN WO9207937-A.

XX XX

XX PD 14-MAY-1992.

XX XX

XX PF 16-OCT-1991; 91WO-FR00810.

XX PR 06-NOV-1990; 90FR-0013731.

XX XX (INRM) INERM INST NAT SANTE RECH MEDICALE.

XX PI Giros B, Martres MP, Schwartz JC, Sokoloff P;

XX XX WPI; 1992-183673/22.

XX DR P-PSDB; AAR24247.

XX XX Human D3 dopaminergic receptor polypeptide(s) - for in-vitro screening of new dopaminergic drugs and diagnosis and treatment of various psychiatric, cardiovascular, neurological and neuro-endocrinological disorders

XX PT

XX PS Claim 4; Page 37; 51pp; French.

XX CC This ORF codes for a mutant D3 receptor with four amino acid substitutions relative to the wild-type receptor sequence (two of the nucleotide substitutions are silent). The mutant receptor retains its dopaminergic receptor properties. The mutant containing all the mutations in the Features Table is preferred but mutants which contain at least one of the mutations are also covered by the invention. See AAQ24803-8.

XX SQ Sequence 1203 BP; 252 A; 351 C; 311 G; 289 T; 0 other;

Query Match 74.8%; Score 17.2; DB 13; Length 1203;

Best Local Similarity 86.4%; Pred. No. 1.9e+02;

Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 CCAACTGGTGGCCATTGACCTT 22

DB 984 CCAATGGTGGCCATTGCTT 1005

RESULT 15

AAA99614

ID AAA99614 standard; cDNA; 1261 BP.

XX AC AAA99614;

XX XX

XX DT 02-FEB-2001 (first entry)

XX XX

us-09-270-437d-12.rng

Mon Apr 21 10:30:08 2003

```

DE Human D2 receptor cDNA.
XX Human; D4 dopamine receptor; cardiovascular system; retinal tissue;
KW vasoregulator; D2 receptor; ss.
XX Homo sapiens.
OS
XX US6121015-A.
XX 19-SEP-2000.
XX 07-JUN-1995; 95US-0475742.
XX 16-JUN-1994; 94US-0261293.
XX 28-JAN-1993; 93US-0014013.
XX (UNIW ) UNIV WASHINGTON.
XX Todd RD, O'Malley KL;
XX WPI; 2000-655527/63.
XX Screening for compounds that selectively bind to a rat D4 dopamine
XX receptor (DDR), useful for identifying dopamine (ant)agonists.
XX comprises exposing cells transfected with a nucleic acid encoding the
XX DDR to candidate compounds -
XX Example 2; Column 35-38; 29pp; English.
XX The present sequence is the human D2 receptor cDNA. The sequence was
XX expressed in the same cell line as the putative rat D4 receptor gene in
XX a method to confirm that the putative D4 receptor does code for a
XX dopamine receptor analogous to the human D4 receptor. The rat D4 receptor
XX was shown to preferentially bind dopamine antagonists such as clozapine.
XX The D4 receptor cDNA is useful for screening drugs which specifically
XX bind to the receptor and have selective effects on the cardiovascular
XX and retinal tissues through interactions with the receptor. Such
XX compounds may act as vasoregulators or may have ionotropic effects. The
XX D4 receptor protein may be used for the production of polyclonal or
XX monoclonal antibodies which recognise the D4 receptor sequence but do not
XX recognise other dopaminergic receptors. The antibodies may be used in
XX immunocytochemical studies and for identification and isolation via flow
XX sorting of D4 expressing cell types.
XX Sequence 1261 BP; 265 A; 363 C; 328 G; 305 T; 0 other;
XX
XX Query Match 74.8%; Score 17.2; DB 21; Length 1261;
XX Best Local Similarity 86.4%; Pred. NO. 1.9e+02;
XX Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
OY 1 CCAACTGGTGGCCATTCAGCTT 22
Db 990 CCAATGGTGGCCATTCAGCTT 1011

```

Search completed: April 18, 2003, 05:45:26
Job time : 105.455 secs

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 18, 2003, 04:53:41 ; Search time 22.4773 Seconds
(without alignments)
313.809 Million cell updates/sec

Title: US-09-270-437D-12

Perfect score: 23
Sequence: 1 ccaactgggtgcccattcagcttc 23

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA: *
1: /cgn2_6/ptodata/1/ina/5A COMB.seq: *
2: /cgn2_6/ptodata/1/ina/5B COMB.seq: *
3: /cgn2_6/ptodata/1/ina/6A COMB.seq: *
4: /cgn2_6/ptodata/1/ina/6B COMB.seq: *
5: /cgn2_6/ptodata/1/ina/PTCUS COMB.seq: *
6: /cgn2_6/ptodata/1/ina/backfiles1.seq: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	23	100.0	1708	4	US-09-061-709-5
C 2	23	100.0	1946	4	US-09-061-709-7
C 3	17.2	74.8	1261	3	US-08-475-742-16
C 4	17.2	74.8	2325	2	US-08-714-677-3
C 5	17.2	74.8	2325	2	US-08-393-540-3
C 6	17.2	74.8	2325	2	US-08-714-537-3
C 7	17.2	74.8	3073	2	US-08-714-677-11
C 8	17.2	74.8	3073	2	US-08-393-540-11
C 9	17.2	74.8	3073	2	US-08-714-537-11
C 10	16.6	72.2	2224	4	US-09-261-855-1
C 11	16.2	70.4	112	1	US-08-518-878B-29
C 12	16.2	70.4	112	1	US-08-518-878B-30
C 13	16.2	70.4	112	1	US-08-294-522B-29
C 14	16.2	70.4	112	1	US-08-294-522B-30
C 15	16.2	70.4	112	2	US-08-807-861A-29
C 16	16.2	70.4	112	2	US-08-807-861A-30
C 17	16.2	70.4	112	2	US-08-470-868A-29
C 18	16.2	70.4	112	2	US-08-470-868A-30
C 19	16.2	70.4	112	3	US-09-210-681-29
C 20	16.2	70.4	112	3	US-09-210-681-30
C 21	16.2	70.4	112	3	US-08-946-719A-29
C 22	16.2	70.4	112	3	US-08-946-719A-30
C 23	16.2	70.4	3941	4	US-09-408-865-2
C 24	16	69.6	1445	1	US-08-324-533-1
C 25	15.8	68.7	1238	4	US-09-313-300-10
C 26	15.8	68.7	1434	4	US-09-129-112-18
C 27	15.6	67.8	401	4	US-09-323-873A-12

c 28	15.6	67.8	453	2	US-08-419-075-5	Sequence 5, Appli
c 29	15.6	67.8	1038	4	US-09-004-838-128	Sequence 128, App
c 30	15.6	67.8	1177	3	US-08-611-587-7	Sequence 7, Appli
c 31	15.6	67.8	1644	2	US-08-458-555-1	Sequence 1, Appli
c 32	15.6	67.8	1719	4	US-09-149-476-152	Sequence 152, App
c 33	15.6	67.8	2277	1	US-08-676-967-5	Sequence 5, Appli
c 34	15.6	67.8	2277	1	US-08-676-974-5	Sequence 5, Appli
c 35	15.6	67.8	2277	2	US-09-098-487-5	Sequence 5, Appli
c 36	15.6	67.8	2299	2	US-09-293-549-13	Sequence 5, Appli
c 37	15.6	67.8	2316	2	US-08-714-677-1	Sequence 13, Appli
c 38	15.6	67.8	2316	2	US-08-393-540-1	Sequence 1, Appli
c 39	15.6	67.8	2316	2	US-08-714-537-1	Sequence 1, Appli
c 40	15.6	67.8	2738	2	US-08-795-868-17	Sequence 17, Appli
c 41	15.6	67.8	2738	4	US-09-303-069-17	Sequence 17, Appli
c 42	15.6	67.8	2738	4	US-09-134-250-17	Sequence 17, Appli
c 43	15.6	67.8	2936	2	US-08-714-677-10	Sequence 10, Appli
c 44	15.6	67.8	2936	2	US-08-393-540-10	Sequence 10, Appli
c 45	15.6	67.8	2936	2	US-08-714-537-10	Sequence 10, Appli

ALIGNMENTS

RESULT 1

US-09-061-709-5/c
; Sequence 5, Application US/09061709B
; Patent No. 6297364
; GENERAL INFORMATION:
; APPLICANT: Chen, Yao-Tseng
; APPLICANT: Gure, Ali
; APPLICANT: Tsang, Solam
; APPLICANT: Stockert, Elisabeth
; APPLICANT: Jager, Elke
; APPLICANT: Knuth, Alexander
; APPLICANT: Old, Lloyd J.
; TITLE OF INVENTION: Isolated Nucleic Acid Molecules Encoding Cancer Associated
; FILE REFERENCE: LUD 5538
; CURRENT APPLICATION NUMBER: US/09/061,709B
; CURRENT FILING DATE: 1998-04-17
; NUMBER OF SEQ ID NOS: 8
; SEQ ID NO 5
; LENGTH: 1708
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
US-09-061-709-5

Query Match 100.0%; Score 23; DB 4; Length 1708;
Best Local Similarity 100.0%; Pred. No. 0.043;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CCAACTGGTGGCCATTGAGTTC 23
Db 68 CCAACTGGTGGCCATTGAGTTC 46

RESULT 2

US-09-061-709-7/c
; Sequence 7, Application US/09061709B
; Patent No. 6297364
; GENERAL INFORMATION:
; APPLICANT: Chen, Yao-Tseng
; APPLICANT: Gure, Ali
; APPLICANT: Tsang, Solam
; APPLICANT: Stockert, Elisabeth
; APPLICANT: Jager, Elke
; APPLICANT: Knuth, Alexander
; APPLICANT: Old, Lloyd J.
; TITLE OF INVENTION: Isolated Nucleic Acid Molecules Encoding Cancer Associated
; FILE REFERENCE: LUD 5538
; CURRENT APPLICATION NUMBER: US/09/061,709B

us-09-270-437d-12.rni

Mon Apr 21 10:30:09 2003

STREET: 419 Seventh Street, N.W., Suite 400
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20004
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/714,677
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/393,540
FILING DATE: 23-FEB-1995
APPLICATION NUMBER: JP 090728
FILING DATE: 06-APR-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 047956
FILING DATE: 23-FEB-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 047940
FILING DATE: 23-FEB-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 090705
FILING DATE: 06-APR-1994
ATTORNEY/AGENT INFORMATION:
NAME: BROWDY, Roger L.
REGISTRATION NUMBER: 25,618
REFERENCE/DOCKET NUMBER: KUBOTA=4
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-628-5197
TELEFAX: 202-737-3528
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 2325 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 1..2325
US-08-714-677-3

Query Match 100.0%; Score 23; DB 4; Length 1946;
Best Local Similarity 100.0%; Pred. No. 0.044; 0; Indels 0; Gaps 0;
Matches 23; Conservative 0; Mismatches 0;

Qy 1 CCAACTGGTGGCCATTGAGCTTC 23
Db 306 CCAACTGGTGGCCATTGAGCTTC 284

RESULT 3
US-08-475-742-16
Sequence 16, Application US/08475742
Patent No. 6121015
GENERAL INFORMATION:
APPLICANT: O'Malley, Karen L
APPLICANT: Todd, Richard D
TITLE OF INVENTION: Gene Encoding the Rat Dopamine D4 Receptor
FILE REFERENCE: WU 102 CON DIV
CURRENT APPLICATION NUMBER: US/08/475,742
CURRENT FILING DATE: 1995-06-07
EARLIER APPLICATION NUMBER: US 08/261,293
EARLIER FILING DATE: 1994-06-16
EARLIER APPLICATION NUMBER: US 08/014,013
EARLIER FILING DATE: 1993-01-28
NUMBER OF SEQ ID NOS: 16
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 16
LENGTH: 1261
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
LOCATION: (1)-(1261)
OTHER INFORMATION: Human d3 cDNA
PUBLICATION INFORMATION:
JOURNAL: C. R. Acad. Sci., D, Sci. Nat.
VOLUME: 311
PAGES: 501-508
DATE: 1990
US-08-475-742-16

Query Match 74.8%; Score 17.2; DB 3; Length 1261;
Best Local Similarity 86.4%; Pred. No. 24; Indels 0; Gaps 0;
Matches 19; Conservative 0; Mismatches 3;

Qy 1 CCAACTGGTGGCCATTGAGCTT 22
Db 990 CCAATGGTGGCCATTGCTT 1011

RESULT 4
US-08-714-677-3
Sequence 3, Application US/08714677
Patent No. 5871977
GENERAL INFORMATION:
APPLICANT: KUBOTA, Michio
APPLICANT: TSUSAKI, Kenji
APPLICANT: MARUTA, Kazuhiko
APPLICANT: SUGIMOTO, Toshiyuki
TITLE OF INVENTION: DNA ENCODING ENZYME, RECOMBINANT DNA AND
TITLE OF INVENTION: ENZYME, TRANSFORMANT, AND THEIR PREPARATIONS AND USES
NUMBER OF SEQUENCES: 17
CORRESPONDENCE ADDRESS:
ADDRESSEE: BROWDY AND NEIMARK, P.L.L.C.

Query Match 74.8%; Score 17.2; DB 2; Length 2325;
Best Local Similarity 86.4%; Pred. No. 26; Indels 0; Gaps 0;
Matches 19; Conservative 0; Mismatches 3;

Qy 2 CCACTGGTGGCCATTGAGCTTC 23
Db 1896 CAACCGGCGCCGCTCAGCTTC 1917

RESULT 5
US-08-393-540-3
Sequence 3, Application US/08393540
Patent No. 5871993
GENERAL INFORMATION:
APPLICANT: KUBOTA, Michio
APPLICANT: TSUSAKI, Kenji
APPLICANT: MARUTA, Kazuhiko
APPLICANT: SUGIMOTO, Toshiyuki
TITLE OF INVENTION: DNA ENCODING ENZYME, RECOMBINANT DNA AND
TITLE OF INVENTION: ENZYME, TRANSFORMANT, AND THEIR PREPARATIONS AND USES
NUMBER OF SEQUENCES: 17
CORRESPONDENCE ADDRESS:
ADDRESSEE: BROWDY AND NEIMARK, P.L.L.C.
STREET: 419 Seventh Street, N.W., Suite 400
CITY: Washington
STATE: D.C.

COUNTRY: USA
ZIP: 20004
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/393,540
FILING DATE: 23-FEB-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 090728
FILING DATE: 06-APR-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 047956
FILING DATE: 23-FEB-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 047940
FILING DATE: 23-FEB-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 090705
FILING DATE: 06-APR-1994
ATTORNEY/AGENT INFORMATION:
NAME: BROWDY, Roger L.
REGISTRATION NUMBER: 25,618
REFERENCE/DOCKET NUMBER: KUBOTA=4
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-628-5197
TELEFAX: 202-737-3528
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 2325 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 1..2325
US-08-393-540-3

Query Match 74.8%; Score 17.2; DB 2; Length 2325;
Best Local Similarity 86.4%; Pred. No. 26;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 2 CAACCTGGTGGCCATTACGCTTC 23
|||||
DB 1896 CAACCGCGCGGCGTTCAGCTTC 1917

RESULT 6
US-08-714-537-3
Sequence 3, Application US/08714537
Patent No. 5871994
GENERAL INFORMATION:
APPLICANT: KUBOTA, Michio
APPLICANT: TSUSAKI, Kenji
APPLICANT: MARUTA, Kazuhiko
APPLICANT: SUGIMOTO, Toshiyuki
TITLE OF INVENTION: DNA ENCODING ENZYME, RECOMBINANT DNA AND
TITLE OF INVENTION: ENZYME, TRANSFORMANT, AND THEIR PREPARATIONS AND USES
NUMBER OF SEQUENCES: 17
CORRESPONDENCE ADDRESS:
ADDRESSEE: BROWDY AND NEIMARK, P.L.L.C.
STREET: 419 Seventh Street, N.W., Suite 400
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20004
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/714,537
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/393,540
FILING DATE: 23-FEB-1995
APPLICATION NUMBER: JP 090728
FILING DATE: 06-APR-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 047956
FILING DATE: 23-FEB-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 047940
FILING DATE: 23-FEB-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 090705
FILING DATE: 06-APR-1994
ATTORNEY/AGENT INFORMATION:
NAME: BROWDY, Roger L.
REGISTRATION NUMBER: 25,618
REFERENCE/DOCKET NUMBER: KUBOTA=4
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-628-5197
TELEFAX: 202-737-3528
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 2325 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 1..2325
US-08-714-537-3

Query Match 74.8%; Score 17.2; DB 2; Length 2325;
Best Local Similarity 86.4%; Pred. No. 26;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 2 CAACCTGGTGGCCATTACGCTTC 23
|||||
DB 1896 CAACCGCGCGGCGTTCAGCTTC 1917

RESULT 7
US-08-714-677-11
Sequence 11, Application US/08714677
Patent No. 5871977
GENERAL INFORMATION:
APPLICANT: KUBOTA, Michio
APPLICANT: TSUSAKI, Kenji
APPLICANT: MARUTA, Kazuhiko
APPLICANT: SUGIMOTO, Toshiyuki
TITLE OF INVENTION: DNA ENCODING ENZYME, RECOMBINANT DNA AND
TITLE OF INVENTION: ENZYME, TRANSFORMANT, AND THEIR PREPARATIONS AND USES
NUMBER OF SEQUENCES: 17
CORRESPONDENCE ADDRESS:
ADDRESSEE: BROWDY AND NEIMARK, P.L.L.C.
STREET: 419 Seventh Street, N.W., Suite 400
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20004
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:

us-09-270-437d-12.rni

Mon Apr 21 10:30:09 2003

```

; APPLICATION NUMBER: US/08/714,677
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA: US/08/393,540
; APPLICATION NUMBER: US/08/393,540
; FILING DATE: 23-FEB-1995
; APPLICATION NUMBER: JP 090728
; FILING DATE: 06-APR-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 047956
; FILING DATE: 23-FEB-1994
; APPLICATION NUMBER: JP 047940
; FILING DATE: 23-FEB-1994
; APPLICATION NUMBER: JP 090705
; FILING DATE: 06-APR-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: BROWDY, Roger L.
; REGISTRATION NUMBER: 25,618
; REFERENCE/DOCKET NUMBER: KUBOTA=4
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-628-5197
; TELEFAX: 202-737-3528
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3073 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cdna
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 678..3002
; US-08-714-677-11

Query Match 74.8%; Score 17.2; DB 2; Length 3073;
Best Local Similarity 86.4%; Pred. No. 27;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 CAACGTGGCCGTCAGCTTC 23
||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2573 CAACCGCGCGCGTTCAGCTTC 2594

RESULT 8
US-08-393-540-11
; Sequence 11, Application US/08393540
; Patent No. 5871993
; GENERAL INFORMATION:
; APPLICANT: KUBOTA, Michio
; APPLICANT: TSUSAKI, Kenji
; APPLICANT: MARUTA, Kazuhiko
; APPLICANT: SUGIMOTO, Toshiyuki
; TITLE OF INVENTION: DNA ENCODING ENZYME, RECOMBINANT DNA AND
; TITLE OF INVENTION: ENZYME, TRANSFORMANT, AND THEIR PREPARATIONS AND USES
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BROWDY AND NEIMARK, P.L.L.C.
; STREET: 419 Seventh Street, N.W., Suite 400
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/393,540
; FILING DATE: 23-FEB-1995
; CLASSIFICATION: 435

```

```

; PRIOR APPLICATION DATA: JP 090728
; APPLICATION NUMBER: JP 090728
; FILING DATE: 06-APR-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 047956
; FILING DATE: 23-FEB-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 047940
; FILING DATE: 23-FEB-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 090705
; FILING DATE: 06-APR-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: BROWDY, Roger L.
; REGISTRATION NUMBER: 25,618
; REFERENCE/DOCKET NUMBER: KUBOTA=4
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-628-5197
; TELEFAX: 202-737-3528
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3073 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cdna
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 678..3002
; US-08-393-540-11

Query Match 74.8%; Score 17.2; DB 2; Length 3073;
Best Local Similarity 86.4%; Pred. No. 27;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 CAACGTGGCCGTCAGCTTC 23
||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2573 CAACCGCGCGCGTTCAGCTTC 2594

RESULT 9
US-08-714-537-11
; Sequence 11, Application US/08714537
; Patent No. 5871994
; GENERAL INFORMATION:
; APPLICANT: KUBOTA, Michio
; APPLICANT: TSUSAKI, Kenji
; APPLICANT: MARUTA, Kazuhiko
; APPLICANT: SUGIMOTO, Toshiyuki
; TITLE OF INVENTION: DNA ENCODING ENZYME, RECOMBINANT DNA AND
; TITLE OF INVENTION: ENZYME, TRANSFORMANT, AND THEIR PREPARATIONS AND USES
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BROWDY AND NEIMARK, P.L.L.C.
; STREET: 419 Seventh Street, N.W., Suite 400
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/714,537
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/393,540
; FILING DATE: 23-FEB-1995
; APPLICATION NUMBER: JP 090728
; FILING DATE: 06-APR-1994

```

;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: JP 047956
;; FILING DATE: 23-FEB-1994
;; PRIOR APPLICATION DATA: JP 047940
;; FILING DATE: 23-FEB-1994
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: JP 090705
;; FILING DATE: 06-APR-1994
;; ATTORNEY/AGENT INFORMATION:
;; NAME: BROWDY, Roger L.
;; REGISTRATION NUMBER: 25,618
;; REFERENCE/DOCKET NUMBER: KUBOTA-4
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 202-628-5197
;; TELEFAX: 202-737-3528
;; INFORMATION FOR SEQ ID NO: 11:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 3073 base pairs
;; TYPE: nucleic acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: cDNA
;; FEATURE:
;; NAME/KEY: CDS
;; LOCATION: 678..3002
US-08-714-537-11

Query Match 74.8%; Score 17.2; DB 2; Length 3073;
Best Local Similarity 86.4%; Pred. No. 27;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 CAACTGGTGGCCATTGAGTTC 23
DB 2573 CAACTGGTGGCCATTGAGTTC 2594

RESULT 10
US-09-261-855-1/c
; Sequence 1, Application US/09261855A
; Patent No. 625055
; GENERAL INFORMATION:
; APPLICANT: Rose, Jeffrey
; TITLE OF INVENTION: THE C-MYC CODING REGION DETERMINANT-BINDING PROTEIN
; FILE REFERENCE: 960296.95131
; CURRENT APPLICATION NUMBER: US/09/261,855A
; CURRENT FILING DATE: 1999-03-03
; NUMBER OF SEQ ID NOS: 46
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 2224
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-261-855-1

Query Match 72.2%; Score 16.6; DB 4; Length 2224;
Best Local Similarity 82.6%; Pred. No. 50;
Matches 19; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 CCAACTGGTGGCCATTGAGTTC 23
DB 563 CCAAGTTGATGGCCATTGAGTTC 541

RESULT 11
US-08-518-8788-29/c
; Sequence 29, Application US/08518878B
; Patent No. 5702902
; GENERAL INFORMATION:
; APPLICANT: Tartaglia, Louis A.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
; TREATMENT OF BODY WEIGHT DISORDERS, INCLUDING OBESITY

;; NUMBER OF SEQUENCES: 57
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Pennie & Edmonds
;; STREET: 1155 Avenue of the Americas
;; CITY: New York
;; STATE: New York
;; COUNTRY: U.S.A.
;; ZIP: 10036-2711
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: PatentIn Release #1.0, Version #1.30
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/518,878B
;; FILING DATE: 23-AUG-1995
;; CLASSIFICATION: 435
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Coruzzi, Laura A.
;; REGISTRATION NUMBER: 30,742
;; REFERENCE/DOCKET NUMBER: 7853-036
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (212) 790-9090
;; TELEFAX: (212) 869-9741/8864
;; TELEX: 66141 PENNIE
;; INFORMATION FOR SEQ ID NO: 29:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 112 base pairs
;; TYPE: nucleic acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
US-08-518-878B-29

Query Match 70.4%; Score 16.2; DB 1; Length 112;
Best Local Similarity 85.7%; Pred. No. 50;
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 3 AACTGGTGGCCATTGAGTTC 23
DB 62 ACCTGAGGGCATTGAGTTC 42

RESULT 12
US-08-518-878B-30/c
; Sequence 30, Application US/08518878B
; Patent No. 5702902
; GENERAL INFORMATION:
; APPLICANT: Tartaglia, Louis A.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
; TREATMENT OF BODY WEIGHT DISORDERS, INCLUDING OBESITY
; NUMBER OF SEQUENCES: 57
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/518,878B
; FILING DATE: 23-AUG-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Coruzzi, Laura A.
; REGISTRATION NUMBER: 30,742
; REFERENCE/DOCKET NUMBER: 7853-036
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090

us-09-270-437d-12.rni

Mon Apr 21 10:30:09 2003

```

; TELEFAX: (212) 869-9741/8864
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 30:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 112 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-518-878B-30

Query Match 70.4%; Score 16.2; DB 1; Length 112;
Best Local Similarity 85.7%; Pred. No. 50;
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 3 AACTGGTGGCCATTGAGTTC 23
Db 62 AGCTGGAGGCATTGAGTTC 42

RESULT 13
US-08-294-522B-29/c
; Sequence 29, Application US/08294522B
; Patent No. 5741666
; GENERAL INFORMATION:
; APPLICANT: Tartaglia, Louis A.
; TITLE OF INVENTION: Compositions and Methods for the
; TREATMENT OF BODY WEIGHT DISORDERS, INCLUDING OBESITY
; NUMBER OF SEQUENCES: 48
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/294,522B
; FILING DATE: 23-AUG-1994
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Coruzzi, Laura A.
; REGISTRATION NUMBER: 30,742
; REFERENCE/DOCKET NUMBER: 7853-015
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-8864/9741
; INFORMATION FOR SEQ ID NO: 30:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 112 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; US-08-294-522B-30

COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/294,522B
; FILING DATE: 23-AUG-1994
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Coruzzi, Laura A.
; REGISTRATION NUMBER: 30,742
; REFERENCE/DOCKET NUMBER: 7853-015
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-8864/9741
; INFORMATION FOR SEQ ID NO: 29:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 112 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; US-08-294-522B-29

Query Match 70.4%; Score 16.2; DB 1; Length 112;
Best Local Similarity 85.7%; Pred. No. 50;
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 3 AACTGGTGGCCATTGAGTTC 23
Db 62 AGCTGGAGGCATTGAGTTC 42

RESULT 14
US-08-294-522B-30/c
; Sequence 30, Application US/08294522B

```

```

; Patent No. 5741666
; GENERAL INFORMATION:
; APPLICANT: Tartaglia, Louis A.
; TITLE OF INVENTION: Compositions and Methods for the
; TREATMENT OF BODY WEIGHT DISORDERS, INCLUDING OBESITY
; NUMBER OF SEQUENCES: 48
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/294,522B
; FILING DATE: 23-AUG-1994
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Coruzzi, Laura A.
; REGISTRATION NUMBER: 30,742
; REFERENCE/DOCKET NUMBER: 7853-015
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-8864/9741
; INFORMATION FOR SEQ ID NO: 30:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 112 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; US-08-294-522B-30

Query Match 70.4%; Score 16.2; DB 1; Length 112;
Best Local Similarity 85.7%; Pred. No. 50;
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 3 AACTGGTGGCCATTGAGTTC 23
Db 62 AGCTGGAGGCATTGAGTTC 42

RESULT 15
US-08-807-861A-29/c
; Sequence 29, Application US/08807861A
; Patent No. 5853975
; GENERAL INFORMATION:
; APPLICANT: Tartaglia, Louis A.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE
; TREATMENT OF BODY WEIGHT DISORDERS, INCLUDING OBESITY
; NUMBER OF SEQUENCES: 64
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/807,861A
; FILING DATE: 26-FEB-1997
; CLASSIFICATION: 514

```



```

; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/518,878
; FILING DATE: 23-AUG-1995
; APPLICATION NUMBER: US 08/470,868
; FILING DATE: 06-JUN-1995
; APPLICATION NUMBER: US 08/294,522
; FILING DATE: 23-AUG-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Coruzzi, Laura A.
; REGISTRATION NUMBER: 30,742
; REFERENCE/DOCKET NUMBER: 7853-066
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-9741/8864
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 29:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 112 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
;
US-08-807-861A-29

```

```

Query Match          70.4%; Score 16.2; DB 2; Length 112;
Best Local Similarity 85.7%; Pred No. 50;
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY  3  ACTGGTGGCCATTTCAGCTTC 23
    |||||
Db   62  AGCTGGAGGGCATTTCAGCTTC 42

```

Search completed: April 18, 2003, 07:34:30
Job time : 23.4773 secs

GenCore version 5.1.4 p5 4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 18, 2003, 06:35:31 ; Search time 95.4848 Seconds
(without alignments)
241.975 Million cell updates/sec

Title: US-09-270-437D-12

Perfect score: 23
Sequence: 1 ccaactgggtggccattcagcttc 23

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 639749 seqs, 502280978 residues

Total number of hits satisfying chosen parameters: 1279498

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- Published Applications NA:*
- 1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
 - 2: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
 - 3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
 - 4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
 - 5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
 - 6: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
 - 7: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
 - 8: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
 - 9: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
 - 10: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
 - 11: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
 - 12: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
 - 13: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
 - 14: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
 - 15: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	23	100.0	1708	10	US-09-899-651-5
C 2	23	100.0	1946	10	US-09-899-651-7
C 3	17.8	77.4	3628	10	US-09-917-800A-1454
C 4	17.2	74.8	1019	7	US-08-781-986A-752
C 5	16.8	73.0	258	10	US-09-864-761-32265
C 6	16.8	73.0	454	10	US-09-864-761-15758
C 7	16.6	72.2	453	10	US-09-960-352-11802
C 8	16.6	72.2	510	10	US-09-864-761-13431
C 9	16.6	72.2	2224	10	US-09-873-637-1
C 10	16.2	70.4	709	10	US-09-864-761-21356
C 11	16.2	70.4	1632	9	US-09-938-842A-1201
C 12	16.2	70.4	1994	10	US-09-864-761-4614
C 13	16.2	70.4	2244	9	US-09-372-348-14
C 14	16.2	70.4	2418	9	US-09-938-842A-1478
C 15	16.2	70.4	8001	9	US-10-261-482-3
C 16	16.2	70.4	536165	9	US-09-939-964-1
C 17	16	69.6	822	10	US-09-893-737-293
C 18	15.8	68.7	553	10	US-09-864-761-7260
C 19	15.8	68.7	1356	10	US-09-815-242-6624

C 20	15.8	68.7	1434	10	US-09-129-112-18	Sequence 18, Appl
C 21	15.8	68.7	1506	10	US-09-822-830A-124	Sequence 124, App
C 22	15.8	68.7	3939	9	US-10-006-856A-226	Sequence 226, App
C 23	15.8	68.7	3939	9	US-10-006-856A-226	Sequence 226, App
C 24	15.8	68.7	3939	9	US-10-015-393A-226	Sequence 226, App
C 25	15.8	68.7	11597	10	US-09-070-927A-222	Sequence 222, App
C 26	15.6	67.8	130	10	US-09-864-761-25655	Sequence 25655, A
C 27	15.6	67.8	157	10	US-09-864-761-24471	Sequence 24471, A
C 28	15.6	67.8	171	10	US-09-864-761-21232	Sequence 21232, A
C 29	15.6	67.8	189	10	US-09-864-761-28773	Sequence 28773, A
C 30	15.6	67.8	256	10	US-09-867-550-2059	Sequence 2059, App
C 31	15.6	67.8	401	9	US-10-011-095-12	Sequence 12, Appl
C 32	15.6	67.8	401	9	US-10-010-667A-12	Sequence 12, Appl
C 33	15.6	67.8	447	10	US-09-864-761-4484	Sequence 4484, App
C 34	15.6	67.8	450	10	US-09-864-761-6525	Sequence 6525, App
C 35	15.6	67.8	455	10	US-09-924-035A-218	Sequence 218, App
C 36	15.6	67.8	457	10	US-09-770-444-511	Sequence 511, App
C 37	15.6	67.8	509	10	US-09-864-761-8998	Sequence 8998, App
C 38	15.6	67.8	579	10	US-09-864-761-7767	Sequence 7767, App
C 39	15.6	67.8	595	10	US-09-920-300A-69	Sequence 69, Appl
C 40	15.6	67.8	626	10	US-09-864-761-12193	Sequence 12193, A
C 41	15.6	67.8	626	12	US-10-033-528-69	Sequence 69, Appl
C 42	15.6	67.8	1038	10	US-09-789-561-17	Sequence 17, Appl
C 43	15.6	67.8	1169	10	US-09-747-835A-11	Sequence 11, Appl
C 44	15.6	67.8	1209	9	US-10-098-841-86	Sequence 86, Appl
C 45	15.6	67.8	1341	10	US-09-954-456-509	Sequence 509, App

ALIGNMENTS

RESULT 1

US-09-899-651-5/c
; Sequence 5, Application US/09899651
; Patent No. US20020111470A1
; GENERAL INFORMATION:
; APPLICANT: Chen, Yao-Tsang
; APPLICANT: Gure, Ali
; APPLICANT: Tsang, Solam
; APPLICANT: Stockert, Elisabeth
; APPLICANT: Jager, Elke
; APPLICANT: Knuth, Alexander
; APPLICANT: Old, Lloyd J.
; TITLE OF INVENTION: Isolated Nucleic Acid Molecules Encoding Cancer
; TITLE OF INVENTION: Associated
; TITLE OF INVENTION: Antigen, The Antigens Per Se, And Uses Thereof
; FILE REFERENCE: LUD 538
; CURRENT APPLICATION NUMBER: US/09/899,651
; CURRENT FILING DATE: 2001-07-06
; PRIOR APPLICATION NUMBER: US/09/061,709
; PRIOR FILING DATE: 1998-04-17
; NUMBER OF SEQ ID NOS: 8
; SEQ ID NO 5
; LENGTH: 1708
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
US-09-899-651-5

Query Match 100.0%; Score 23; DB 10; Length 1708;
Best Local Similarity 100.0%; Pred. No. 0.065;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 CCAACTGGTGGCCATTGAGCTTC 23
Db 68 CCAACTGGTGGCCATTGAGCTTC 46

RESULT 2

US-09-899-651-7/c
; Sequence 7, Application US/09899651
; Patent No. US20020111470A1
; GENERAL INFORMATION:

Mon Apr 21 10:30:11 2003

APPLICANT: Chen, Yao-Tseung
 APPLICANT: Gure, Ali
 APPLICANT: Tsang, Solam
 APPLICANT: Stockert, Elisabeth
 APPLICANT: Jager, Elke
 APPLICANT: Knuth, Alexander
 APPLICANT: Old, Lloyd J.
 TITLE OF INVENTION: Isolated Nucleic Acid Molecules Encoding Cancer
 TITLE OF INVENTION: Associated
 TITLE OF INVENTION: Antigen, The Antigens Per Se, And Uses Thereof
 FILE REFERENCE: LUD 5538
 CURRENT APPLICATION NUMBER: US/09/899,651
 CURRENT FILING DATE: 2001-07-06
 PRIOR APPLICATION NUMBER: US/09/061,709
 PRIOR FILING DATE: 1998-04-17
 NUMBER OF SEQ ID NOS: 8
 SEQ ID NO 7
 LENGTH: 1946
 TYPE: DNA
 ORGANISM: Homo sapiens
 FEATURE:
 US-09-899-651-7

Query Match 100.0%; Score 23; DB 10; Length 1946;
 Best Local Similarity 100.0%; Pred. No. 0.065;
 Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCACTGGTGGCCATTCAGCTTC 23
 DB 306 CCACTGGTGGCCATTCAGCTTC 284

RESULT 3
 US-09-917-800A-1454/c
 ; Sequence 1454, Application US/09917800A
 ; Patent No. US20020119462A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Mendrick, Donna
 ; APPLICANT: Porter, Mark
 ; APPLICANT: Johnson, Kory
 ; APPLICANT: Castle, Arthur
 ; APPLICANT: Elashoff, Michael
 ; APPLICANT: Gene Logic, Inc.
 ; TITLE OF INVENTION: Molecular Toxicology Modeling
 ; FILE REFERENCE: 44921-5038-US
 ; CURRENT APPLICATION NUMBER: US/09/917,800A
 ; CURRENT FILING DATE: 2001-07-31
 ; PRIOR APPLICATION NUMBER: US 60/222,040
 ; PRIOR FILING DATE: 2000-07-31
 ; PRIOR APPLICATION NUMBER: US 60/222,880
 ; PRIOR FILING DATE: 2000-11-02
 ; PRIOR APPLICATION NUMBER: US 60/290,029
 ; PRIOR FILING DATE: 2001-05-11
 ; PRIOR APPLICATION NUMBER: US 60/290,645
 ; PRIOR FILING DATE: 2001-05-15
 ; PRIOR APPLICATION NUMBER: US 60/292,336
 ; PRIOR FILING DATE: 2001-05-22
 ; PRIOR APPLICATION NUMBER: US 60/295,798
 ; PRIOR FILING DATE: 2001-06-06
 ; PRIOR APPLICATION NUMBER: US 60/297,457
 ; PRIOR FILING DATE: 2001-06-13
 ; PRIOR APPLICATION NUMBER: US 60/298,884
 ; PRIOR FILING DATE: 2001-06-19
 ; PRIOR APPLICATION NUMBER: US 60/303,459
 ; PRIOR FILING DATE: 2001-07-09
 ; NUMBER OF SEQ ID NOS: 1740
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 1454
 ; LENGTH: 3628
 ; TYPE: DNA
 ; ORGANISM: Rattus norvegicus
 ; FEATURE:
 ; OTHER INFORMATION: Genbank Accession No. US20020119462A1 U20796

US-09-917-800A-1454

Query Match 77.4%; Score 17.8; DB 10; Length 3628;
 Best Local Similarity 90.5%; Pred. No. 25;
 Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 AACTGGTGGCCATTCAGCTTC 23
 DB 1355 AGCTGGAGGCCATTCAGCTTC 1335

RESULT 4
 US-08-781-986A-752
 ; Sequence 752, Application US/08781986A
 ; Publication No. US20030054436A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Charles Kunsch
 ; TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
 ; NUMBER OF SEQUENCES: 5255
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Human Genome Sciences, Inc.
 ; STREET: 9410 Key West Avenue
 ; CITY: Rockville
 ; STATE: Maryland
 ; COUNTRY: USA
 ; ZIP: 20850
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
 ; COMPUTER: HP Vectra 486/33
 ; OPERATING SYSTEM: MSDOS version 6.2
 ; SOFTWARE: ASCII Text
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/781,986A
 ; FILING DATE:
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER:
 ; FILING DATE:
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Benson, Bob
 ; REGISTRATION NUMBER: 30,446
 ; REFERENCE/DOCKET NUMBER: PB248PP
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (301) 309-8504
 ; TELEFAX: (301) 309-8512
 ; INFORMATION FOR SEQ ID NO: 752:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 1019 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: double
 ; TOPOLOGY: linear
 ; US-08-781-986A-752

Query Match 74.8%; Score 17.2; DB 7; Length 1019;
 Best Local Similarity 86.4%; Pred. No. 43;
 Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 CAACTGGTGGCCATTCAGCTTC 23
 DB 640 CAACTGGTGGTAATTCAGCTTC 661

RESULT 5
 US-09-864-761-32265/c
 ; Sequence 32265, Application US/09864761
 ; Patent No. US20020048763A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Penn, Sharron G.
 ; APPLICANT: Rank, David R.
 ; APPLICANT: Hanzel, David K.
 ; APPLICANT: Chen, Wensheng
 ; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
 ; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY

FILE REFERENCE: Aemica-X-1
CURRENT APPLICATION NUMBER: US/09/864,761
CURRENT FILING DATE: 2001-05-23
PRIOR APPLICATION NUMBER: US 60/180,312
PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: US 60/207,456
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: US 09/632,366
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: GB 24263.6
PRIOR FILING DATE: 2000-10-04
PRIOR APPLICATION NUMBER: US 60/236,359
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: PCT/US01/00666
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00667
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00664
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00669
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00665
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00668
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00663
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00662
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00661
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00670
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: US 60/234,687
PRIOR FILING DATE: 2000-09-21
PRIOR APPLICATION NUMBER: US 09/608,408
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: US 09/774,203
PRIOR FILING DATE: 2001-01-29
NUMBER OF SEQ ID NOS: 49117
SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
SEQ ID NO 32265
LENGTH: 258
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AC002093.1
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.5
OTHER INFORMATION: EST HUMAN HIT: BE798126.1, EVALUE 1.00e-114
OTHER INFORMATION: SWISSPROT HIT: P34369, EVALUE 8.00e-14
OTHER INFORMATION: NT HIT: G11436714, EVALUE 1.00e-121
US-09-864-761-32265

Query Match 73.0%; Score 16.8; DB 10; Length 258;
Best Local Similarity 90.0%; Pred. No. 58;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy 2 CAACGTGGTGGCCATTCAGCT 21
Db 95 CCACTGGAGGCCATTCAGCT 76

RESULT 6
US-09-864-761-15759/c
Sequence 15758, Application US/09864761
Patent No. US20020048763A1
GENERAL INFORMATION:
APPLICANT: Penn, Sharon G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
APPLICANT: Chen, Wensheng
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY

FILE REFERENCE: Aemica-X-1
CURRENT APPLICATION NUMBER: US/09/864,761
CURRENT FILING DATE: 2001-05-23
PRIOR APPLICATION NUMBER: US 60/180,312
PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: US 60/207,456
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: US 09/632,366
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: GB 24263.6
PRIOR FILING DATE: 2000-10-04
PRIOR APPLICATION NUMBER: US 60/236,359
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: PCT/US01/00666
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00667
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00664
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00669
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00665
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00668
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00663
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00662
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00661
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00670
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: US 60/234,687
PRIOR FILING DATE: 2000-09-21
PRIOR APPLICATION NUMBER: US 09/608,408
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: US 09/774,203
PRIOR FILING DATE: 2001-01-29
NUMBER OF SEQ ID NOS: 49117
SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
SEQ ID NO 15758
LENGTH: 454
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AC002093.1
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.5
NAME/KEY: unsure
LOCATION: 256
US-09-864-761-15758

Query Match 73.0%; Score 16.8; DB 10; Length 454;
Best Local Similarity 90.0%; Pred. No. 62;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy 2 CAACGTGGTGGCCATTCAGCT 21
Db 376 CCACTGGAGGCCATTCAGCT 357

RESULT 7
US-09-960-352-11802/c
Sequence 11802, Application US/09960352
Patent No. US20020137139A1
GENERAL INFORMATION:
APPLICANT: Warren, Wesley C.
APPLICANT: Rao, Nengbing
APPLICANT: Byatt, John C.
APPLICANT: Mathialagan, Nagappan
TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
TITLE OF INVENTION: MUSCLE AND FAT DEPOSITION
FILE REFERENCE: 16511.006/37-21(10298)C

us-09-270-437d-12.rnpb

Mon Apr 21 10:30:11 2003

```
;
;
; CURRENT APPLICATION NUMBER: US/09/960,352
; CURRENT FILING DATE: 2001-09-24
; NUMBER OF SEQ ID NOS: 15112
; SEQ ID NO 11802
; LENGTH: 453
; TYPE: DNA
; ORGANISM: Bos taurus
; OTHER INFORMATION: Clone ID: 50-LIB34-078-Q1-E1-E10
US-09-960-352-11802

Query Match      72.2%; Score 16.6; DB 10; Length 453;
Best Local Similarity 82.6%; Pred. No. 77;
Matches 19; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 CCAACTGGTGGCCATTGAGCTTC 23
Db 331 CCAACTGGTGGCCAGTCCAGCTCC 309

RESULT 8
US-09-864-761-13431/c
; Sequence 13431, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aemica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annonax Sequence Listing Engine vers. 1.1
; SEQ ID NO 13431
; LENGTH: 510

;
;
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC023314.3
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 2.1
US-09-864-761-13431

Query Match      72.2%; Score 16.6; DB 10; Length 510;
Best Local Similarity 82.6%; Pred. No. 78;
Matches 19; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 CCAACTGGTGGCCATTGAGCTTC 23
Db 351 CCAACTGGTGGCCATTGAGCTTC 329

RESULT 9
US-09-873-637-1/c
; Sequence 1, Application US/09873637
; Patent No. US20020061543A1
; GENERAL INFORMATION:
; APPLICANT: Ross, Jeffrey
; TITLE OF INVENTION: THE C-MYC CODING REGION DETERMINANT-BINDING PROTEIN
; TITLE OF INVENTION: (CRD-BP) AND ITS NUCLEIC ACID SEQUENCE
; FILE REFERENCE: 960296.95131
; CURRENT APPLICATION NUMBER: US/09/873,637
; CURRENT FILING DATE: 2001-06-04
; NUMBER OF SEQ ID NOS: 46
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 2224
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-873-637-1

Query Match      72.2%; Score 16.6; DB 10; Length 2224;
Best Local Similarity 82.6%; Pred. No. 91;
Matches 19; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 CCAACTGGTGGCCATTGAGCTTC 23
Db 563 CCAGTTGATGGCCATTGAGCTTC 541

RESULT 10
US-09-864-761-21356
; Sequence 21356, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
; FILE REFERENCE: Aemica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annonax Sequence Listing Engine vers. 1.1
; SEQ ID NO 13431
; LENGTH: 510
```


us-09-270-437d-12.rnbp

Mon Apr 21 10:30:11 2003

```
; OTHER INFORMATION: MAP TO AC004230.1
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 5.3
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 68
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 5.9
; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 5.8
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 12
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 51
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 45
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 9.2
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 5.6
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 37
US-09-864-761-4614

Query Match          70.4%; Score 16.2; DB 10; Length 1994;
Best Local Similarity 85.7%; Pred. No. 1.4e+02;
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 3 CCACTGGTGGCCATTTCAGCTTC 23
Db 319 AACTGGGGCCCTTCAGCTTC 339

RESULT 13
US-09-372-348-14/c
; Sequence 14, Application US/09372348A
; Publication No. US20030028005A1
; GENERAL INFORMATION:
; APPLICANT: Hedrick, Joseph A.
; APPLICANT: Wang, Luquan
; APPLICANT: Zlotnik, Albert
; APPLICANT: Murgolo, Nicholas R.
; APPLICANT: Greene, Jonathan R.
; APPLICANT: Johnston, James A.
; APPLICANT: Bazan, J. Fernando
; TITLE OF INVENTION: Mammalian Proteins; Related Reagents and Methods
; FILE REFERENCE: DX0884K(2d)
; CURRENT APPLICATION NUMBER: US/09/372,348A
; CURRENT FILING DATE: 1999-08-11
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 14
; LENGTH: 2244
; TYPE: DNA
; ORGANISM: primate
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(2244)
US-09-372-348-14

Query Match          70.4%; Score 16.2; DB 9; Length 2244;
Best Local Similarity 85.7%; Pred. No. 1.4e+02;
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 CCACTGGTGGCCATTTCAGCTTC 21
Db 2047 CCACCTGCTGGCCATTTCACCT 2027

RESULT 14
US-09-938-842A-1478
; Sequence 1478, Application US/09938842A
; Patent No. US20020160378A1
; GENERAL INFORMATION:
; APPLICANT: Harper, Jeff
; APPLICANT: Kreps, Joel
; APPLICANT: Wang, Xun
; APPLICANT: Zhu, Tong
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
; FILE REFERENCE: SCRIPI300-3
; CURRENT APPLICATION NUMBER: US/09/938,842A
; CURRENT FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: US 60/227,866

; OTHER INFORMATION:
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 68
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 5.9
; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 5.8
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 12
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 51
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 45
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 9.2
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 5.6
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 37
US-09-864-761-4614

Query Match          70.4%; Score 16.2; DB 9; Length 2418;
Best Local Similarity 85.7%; Pred. No. 1.4e+02;
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 CCACTGGTGGCCATTTCAGCTTC 21
Db 46 CCAACTTGTGGCTATGCAGCT 66

RESULT 15
US-10-261-482-3/c
; Sequence 3, Application US/10261482
; Publication No. US20030036089A1
; GENERAL INFORMATION:
; APPLICANT: Wei, Ming-Hui et al
; TITLE OF INVENTION: ISOLATED HUMAN G-PROTEIN COUPLED
; TITLE OF INVENTION: RECEPTORS, NUCLEIC ACID MOLECULES ENCODING HUMAN GPCR
; TITLE OF INVENTION: PROTEINS, AND USES THEREOF
; FILE REFERENCE: CL000869CON
; CURRENT APPLICATION NUMBER: US/10/261,482
; CURRENT FILING DATE: 2002-10-02
; PRIOR APPLICATION NUMBER: 09/684,393
; PRIOR FILING DATE: 2000-10-10
; PRIOR APPLICATION NUMBER: 60/172,600
; PRIOR FILING DATE: 1999-12-20
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 8001
; TYPE: DNA
; ORGANISM: Human
US-10-261-482-3

Query Match          70.4%; Score 16.2; DB 9; Length 8001;
Best Local Similarity 85.7%; Pred. No. 1.6e+02;
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 3 AACTGGTGGCCATTTCAGCTTC 23
Db 1181 AACTGGGGGCACTCAGCTTC 1161

Search completed: April 18, 2003, 10:14:07
Job time : 97.6515 secs
```


GenCore version 5.1.4 p5 4578
Copyright (C) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 18, 2003, 05:48:17 ; Search time 1266.05 Seconds
(without alignments)
456.759 Million cell updates/sec

Title: US-09-270-437D-12

Perfect score: 23

Sequence: 1 ccaactgggtggtccattcagcttc 23

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 24791104 seqs, 12571243825 residues

Total number of hits satisfying chosen parameters: 49582208

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

- Database :
- Pending Patents NA Main:*
- 1: /cgn2_6/ptodata/1/pna/US06 COMB.seq:*
 - 2: /cgn2_6/ptodata/1/pna/US06 COMB.seq:*
 - 3: /cgn2_6/ptodata/1/pna/US07 COMB.seq:*
 - 4: /cgn2_6/ptodata/1/pna/US08 COMB.seq:*
 - 5: /cgn2_6/ptodata/1/pna/US08 COMB.seq:*
 - 6: /cgn2_6/ptodata/1/pna/US08 COMB.seq:*
 - 7: /cgn2_6/ptodata/1/pna/US08 COMB.seq:*
 - 8: /cgn2_6/ptodata/1/pna/US08 COMB.seq:*
 - 9: /cgn2_6/ptodata/1/pna/US08 COMB.seq:*
 - 10: /cgn2_6/ptodata/1/pna/US08 COMB.seq:*
 - 11: /cgn2_6/ptodata/1/pna/US08 COMB.seq:*
 - 12: /cgn2_6/ptodata/1/pna/US08 COMB.seq:*
 - 13: /cgn2_6/ptodata/1/pna/US08 COMB.seq:*
 - 14: /cgn2_6/ptodata/1/pna/US09 COMB.seq:*
 - 15: /cgn2_6/ptodata/1/pna/US09 COMB.seq:*
 - 16: /cgn2_6/ptodata/1/pna/US09 COMB.seq:*
 - 17: /cgn2_6/ptodata/1/pna/US09 COMB.seq:*
 - 18: /cgn2_6/ptodata/1/pna/US09 COMB.seq:*
 - 19: /cgn2_6/ptodata/1/pna/US09 COMB.seq:*
 - 20: /cgn2_6/ptodata/1/pna/US09 COMB.seq:*
 - 21: /cgn2_6/ptodata/1/pna/US09 COMB.seq:*
 - 22: /cgn2_6/ptodata/1/pna/US09 COMB.seq:*
 - 23: /cgn2_6/ptodata/1/pna/US09 COMB.seq:*
 - 24: /cgn2_6/ptodata/1/pna/US09 COMB.seq:*
 - 25: /cgn2_6/ptodata/1/pna/US09 COMB.seq:*
 - 26: /cgn2_6/ptodata/1/pna/US09 COMB.seq:*
 - 27: /cgn2_6/ptodata/1/pna/US09 COMB.seq:*
 - 28: /cgn2_6/ptodata/1/pna/US09 COMB.seq:*
 - 29: /cgn2_6/ptodata/1/pna/US09 COMB.seq:*
 - 30: /cgn2_6/ptodata/1/pna/US09 COMB.seq:*
 - 31: /cgn2_6/ptodata/1/pna/US09 COMB.seq:*
 - 32: /cgn2_6/ptodata/1/pna/US09 COMB.seq:*
 - 33: /cgn2_6/ptodata/1/pna/US09 COMB.seq:*
 - 34: /cgn2_6/ptodata/1/pna/US09 COMB.seq:*
 - 35: /cgn2_6/ptodata/1/pna/US09 COMB.seq:*
 - 36: /cgn2_6/ptodata/1/pna/US09 COMB.seq:*
 - 37: /cgn2_6/ptodata/1/pna/US09 COMB.seq:*
 - 38: /cgn2_6/ptodata/1/pna/US09 COMB.seq:*
 - 39: /cgn2_6/ptodata/1/pna/US09 COMB.seq:*
 - 40: /cgn2_6/ptodata/1/pna/US09 COMB.seq:*
 - 41: /cgn2_6/ptodata/1/pna/US09 COMB.seq:*
 - 42: /cgn2_6/ptodata/1/pna/US09 COMB.seq:*
 - 43: /cgn2_6/ptodata/1/pna/US09 COMB.seq:*

- 44: /cgn2_6/ptodata/1/pna/US6000 COMB.seq:*
- 45: /cgn2_6/ptodata/1/pna/US6001 COMB.seq:*
- 46: /cgn2_6/ptodata/1/pna/US6002 COMB.seq:*
- 47: /cgn2_6/ptodata/1/pna/US6003 COMB.seq:*
- 48: /cgn2_6/ptodata/1/pna/US6004 COMB.seq:*
- 49: /cgn2_6/ptodata/1/pna/US6005 COMB.seq:*
- 50: /cgn2_6/ptodata/1/pna/US6006 COMB.seq:*
- 51: /cgn2_6/ptodata/1/pna/US6007 COMB.seq:*
- 52: /cgn2_6/ptodata/1/pna/US6008 COMB.seq:*
- 53: /cgn2_6/ptodata/1/pna/US6009 COMB.seq:*
- 54: /cgn2_6/ptodata/1/pna/US6010 COMB.seq:*
- 55: /cgn2_6/ptodata/1/pna/US6011 COMB.seq:*
- 56: /cgn2_6/ptodata/1/pna/US6012 COMB.seq:*
- 57: /cgn2_6/ptodata/1/pna/US6013 COMB.seq:*
- 58: /cgn2_6/ptodata/1/pna/US6014 COMB.seq:*
- 59: /cgn2_6/ptodata/1/pna/US6015 COMB.seq:*
- 60: /cgn2_6/ptodata/1/pna/US6016 COMB.seq:*
- 61: /cgn2_6/ptodata/1/pna/US6017 COMB.seq:*
- 62: /cgn2_6/ptodata/1/pna/US6018 COMB.seq:*
- 63: /cgn2_6/ptodata/1/pna/US6019 COMB.seq:*
- 64: /cgn2_6/ptodata/1/pna/US6020 COMB.seq:*
- 65: /cgn2_6/ptodata/1/pna/US6021 COMB.seq:*
- 66: /cgn2_6/ptodata/1/pna/US6022 COMB.seq:*
- 67: /cgn2_6/ptodata/1/pna/US6023 COMB.seq:*
- 68: /cgn2_6/ptodata/1/pna/US6024 COMB.seq:*
- 69: /cgn2_6/ptodata/1/pna/US6025 COMB.seq:*
- 70: /cgn2_6/ptodata/1/pna/US6026 COMB.seq:*
- 71: /cgn2_6/ptodata/1/pna/US6027 COMB.seq:*
- 72: /cgn2_6/ptodata/1/pna/US6028 COMB.seq:*
- 73: /cgn2_6/ptodata/1/pna/US6029 COMB.seq:*
- 74: /cgn2_6/ptodata/1/pna/US6030 COMB.seq:*
- 75: /cgn2_6/ptodata/1/pna/US6031 COMB.seq:*
- 76: /cgn2_6/ptodata/1/pna/US6032 COMB.seq:*
- 77: /cgn2_6/ptodata/1/pna/US6033 COMB.seq:*
- 78: /cgn2_6/ptodata/1/pna/US6034 COMB.seq:*
- 79: /cgn2_6/ptodata/1/pna/US6035 COMB.seq:*
- 80: /cgn2_6/ptodata/1/pna/US6036 COMB.seq:*
- 81: /cgn2_6/ptodata/1/pna/US6037 COMB.seq:*
- 82: /cgn2_6/ptodata/1/pna/US6038 COMB.seq:*
- 83: /cgn2_6/ptodata/1/pna/US6039 COMB.seq:*
- 84: /cgn2_6/ptodata/1/pna/US6040 COMB.seq:*
- 85: /cgn2_6/ptodata/1/pna/US6041 COMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
C 1	23	100.0	482	56	US-60-126-246-152	Sequence 152, App
C 2	23	100.0	1708	16	US-09-270-437-5	Sequence 5, Appli
C 3	23	100.0	1708	33	US-09-899-651-5	Sequence 5, Appli
C 4	23	100.0	1738	24	US-09-625-102-2082	Sequence 2082, Ap
C 5	23	100.0	1946	16	US-09-270-437-7	Sequence 7, Appli
C 6	23	100.0	1946	33	US-09-899-651-7	Sequence 7, Appli
C 7	23	100.0	2512	23	US-09-617-182-2240	Sequence 2240, Ap
C 8	19.8	86.1	412	18	US-09-487-566A-4721	Sequence 4721, Ap
C 9	19.8	86.1	412	18	US-09-487-566A-4721	Sequence 4721, Ap
C 10	19.4	84.3	235	11	US-08-792-586-517	Sequence 517, App
C 11	19.4	84.3	235	20	US-09-534-843-43302	Sequence 43302, A
C 12	18.8	81.7	246	15	US-09-107-426-998	Sequence 998, App
C 13	18.8	81.7	246	49	US-60-051-750-998	Sequence 998, App
C 14	18.8	81.7	328	29	US-09-726-807-731	Sequence 731, App
C 15	18.8	81.7	338	8	US-08-401-791A-10518	Sequence 10518, A
C 16	18.8	81.7	338	8	US-08-401-791B-10518	Sequence 10518, A
C 17	18.8	81.7	348	25	US-09-652-121-3346	Sequence 3346, Ap
C 18	18.8	81.7	367	16	US-09-270-767-1742	Sequence 1742, Ap
C 19	18.8	81.7	367	16	US-09-270-767-17024	Sequence 17024, A
C 20	18.8	81.7	367	16	US-09-270-849B-12212	Sequence 12212, A
C 21	18.8	81.7	415	1	PCT-US01-01354-3063	Sequence 3063, Ap

us-09-270-437d-12.rnmpm

Mon Apr 21 10:30:15 2003

Sequence 3063, Ap
Sequence 1063, Ap
Sequence 14442, A
Sequence 213, App
Sequence 14442, A
Sequence 261, App
Sequence 5712, Ap
Sequence 205, App
Sequence 205, App
Sequence 1129, Ap
Sequence 1129, Ap
Sequence 6485, Ap
Sequence 3047, Ap
Sequence 282, App
Sequence 348, App
Sequence 1243, Ap
Sequence 298, App
Sequence 298, App
Sequence 329, App
Sequence 49, App1
Sequence 28, App1
Sequence 11836, A
Sequence 40121, A
Sequence 39753, A

ALIGNMENTS

RESULT 1
US-60-126-246-152/c
; Sequence 152, Application US/60126246
; GENERAL INFORMATION:
; APPLICANT: Hodgson, David M.
; APPLICANT: Lincoln, Stephen L.
; APPLICANT: Russo, Frank D.
; APPLICANT: Spiro, Peter A.
; APPLICANT: Brantner, Shawn R.
; APPLICANT: Dufour, Gerard J.
; APPLICANT: Cohen, Howard J.
; APPLICANT: Rosen, Bruce
; APPLICANT: Shah, Purvi
; APPLICANT: Chalap, Michael S.
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Jones, Anissa L.
; APPLICANT: Yu, Jimmy Y.
; APPLICANT: Greenawalt, Lila B.
; APPLICANT: Panzer, Scott R.
; APPLICANT: Roseberry, Ana M.
; APPLICANT: Wright, Rachel J.
; TITLE OF INVENTION: MOLECULES ASSOCIATED WITH GROWTH AND DEVELOPMENT
; FILE REFERENCE: PT-0002 P
; CURRENT APPLICATION NUMBER: US/60/126,246
; CURRENT FILING DATE: 1999-03-24
; NUMBER OF SEQ ID NOS: 413
; SOFTWARE: PERL Program
; SEQ ID NO 152
; LENGTH: 482
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: 39667.1
US-60-126-246-152

Query Match 100.0%; Score 23; DB 56; Length 482;
Best Local Similarity 100.0%; Pred. No. 2;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCAACTGGTGGCCATTCAGCTTC 23
DB 42 CCAACTGGTGGCCATTCAGCTTC 20

RESULT 2
US-09-270-437-5/c
; Sequence 5, Application US/09270437A
; GENERAL INFORMATION:
; APPLICANT: Chen, Yao-Tsang
; APPLICANT: Gure, Ali
; APPLICANT: Tsang, Solam
; APPLICANT: Stockert, Elisabeth
; APPLICANT: Jager, Elke
; APPLICANT: Knuth, Alexander
; APPLICANT: Old, Lloyd J.
; TITLE OF INVENTION: Isolated Nucleic Acid Molecules Encoding Cancer Associated Antigen
; TITLE OF INVENTION: Antigens Per Se, And Uses Thereof
; FILE REFERENCE: LUD 5538.1 PCT
; CURRENT APPLICATION NUMBER: US/09/270,437A
; CURRENT FILING DATE: 1999-03-16
; NUMBER OF SEQ ID NOS: 8
; SEQ ID NO 5
; LENGTH: 1708
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
US-09-270-437-5

Query Match 100.0%; Score 23; DB 16; Length 1708;
Best Local Similarity 100.0%; Pred. No. 2.5;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCAACTGGTGGCCATTCAGCTTC 23
DB 68 CCAACTGGTGGCCATTCAGCTTC 46

RESULT 3
US-09-899-651-5/c
; Sequence 5, Application US/09899651
; GENERAL INFORMATION:
; APPLICANT: Chen, Yao-Tsang
; APPLICANT: Gure, Ali
; APPLICANT: Tsang, Solam
; APPLICANT: Stockert, Elisabeth
; APPLICANT: Jager, Elke
; APPLICANT: Knuth, Alexander
; APPLICANT: Old, Lloyd J.
; TITLE OF INVENTION: Isolated Nucleic Acid Molecules Encoding Cancer
; TITLE OF INVENTION: Associated
; TITLE OF INVENTION: Antigen, The Antigens Per Se, And Uses Thereof
; FILE REFERENCE: LUD 5538
; CURRENT APPLICATION NUMBER: US/09/899,651
; CURRENT FILING DATE: 2001-07-06
; PRIOR APPLICATION NUMBER: US/09/061,709
; PRIOR FILING DATE: 1998-04-17
; NUMBER OF SEQ ID NOS: 8
; SEQ ID NO 5
; LENGTH: 1708
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
US-09-899-651-5

Query Match 100.0%; Score 23; DB 33; Length 1708;
Best Local Similarity 100.0%; Pred. No. 2.5;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCAACTGGTGGCCATTCAGCTTC 23
DB 68 CCAACTGGTGGCCATTCAGCTTC 46

RESULT 4
US-09-625-102-2082
; Sequence 2082, Application US/09625102

```
/ GENERAL INFORMATION:
/ APPLICANT: Loring, Jeanne F.
/ FILE REFERENCE: LUD 5538
/ TITLE OF INVENTION: MARKERS OF NEURONAL DIFFERENTIATION AND MORPHOGENESIS
/ CURRENT APPLICATION NUMBER: PA-0026 US
/ CURRENT FILING DATE: 2000-07-25
/ NUMBER OF SEQ ID NOS: 2217
/ SOFTWARE: PERL Program
/ SEQ ID NO 2082
/ LENGTH: 1738
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ NAME/KEY: misc feature
/ OTHER INFORMATION: Incyte ID No: 457674.1
US-09-625-102-2082

Query Match      100.0%; Score 23; DB 24; Length 1738;
Best Local Similarity 100.0%; Pred. No. 2.5;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CCAACTGGTGGCCATTGAGCTTC 23
Db 1697 CCAACTGGTGGCCATTGAGCTTC 1719

RESULT 5
US-09-270-437-7/c
/ Sequence 7, Application US/09270437A
/ GENERAL INFORMATION:
/ APPLICANT: Chen, Yao-Tsang
/ APPLICANT: Gure, Ali
/ APPLICANT: Teang, Solam
/ APPLICANT: Stockert, Elisabeth
/ APPLICANT: Jager, Elke
/ APPLICANT: Knuth, Alexander
/ APPLICANT: Old, Lloyd J.
/ TITLE OF INVENTION: Isolated Nucleic Acid Molecules Encoding Cancer Associated Antigens
/ FILE REFERENCE: LUD 5538.1 PCT
/ CURRENT APPLICATION NUMBER: US/09/270.437A
/ CURRENT FILING DATE: 1999-03-16
/ NUMBER OF SEQ ID NOS: 8
/ SEQ ID NO 7
/ LENGTH: 1946
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ NAME/KEY: misc feature
/ OTHER INFORMATION: Isolated Nucleic Acid Molecules Encoding Cancer Associated Antigens

US-09-270-437-7

Query Match      100.0%; Score 23; DB 16; Length 1946;
Best Local Similarity 100.0%; Pred. No. 2.5;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CCAACTGGTGGCCATTGAGCTTC 23
Db 306 CCAACTGGTGGCCATTGAGCTTC 284

RESULT 6
US-09-899-651-7/c
/ Sequence 7, Application US/09899651
/ GENERAL INFORMATION:
/ APPLICANT: Chen, Yao-Tsang
/ APPLICANT: Gure, Ali
/ APPLICANT: Teang, Solam
/ APPLICANT: Stockert, Elisabeth
/ APPLICANT: Jager, Elke
/ APPLICANT: Knuth, Alexander
/ APPLICANT: Old, Lloyd J.
/ TITLE OF INVENTION: Isolated Nucleic Acid Molecules Encoding Cancer Associated Antigens

US-09-899-651-7/c

Query Match      100.0%; Score 23; DB 23; Length 1946;
Best Local Similarity 100.0%; Pred. No. 2.5;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CCAACTGGTGGCCATTGAGCTTC 23
Db 747 CCAACTGGTGGCCATTGAGCTTC 725

RESULT 8
US-09-487-566-4721/c
/ Sequence 4721, Application US/09487566
/ GENERAL INFORMATION:
/ APPLICANT: Rosen, et. al.
/ TITLE OF INVENTION: Human Genes, Sequences, and Expression Products
/ FILE REFERENCE: P044
/ CURRENT APPLICATION NUMBER: US/09/487,566
/ CURRENT FILING DATE: 2000-01-19
/ NUMBER OF SEQ ID NOS: 5506
/ SOFTWARE: PatentIn Ver. 2.0
/ SEQ ID NO 4721
/ LENGTH: 412
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ NAME/KEY: misc feature
/ LOCATION: (163)
/ OTHER INFORMATION: n equals a,t,g, or c
/ FEATURE:
```

us-09-270-437d-12.rnnp

Mon Apr 21 10:30:15 2003

```

NAME/KEY: misc feature
LOCATION: (282)
OTHER INFORMATION: n equals a,t,g, or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (313)
OTHER INFORMATION: n equals a,t,g, or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (387)
OTHER INFORMATION: n equals a,t,g, or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (395)
OTHER INFORMATION: n equals a,t,g, or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (402)
OTHER INFORMATION: n equals a,t,g, or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (412)
OTHER INFORMATION: n equals a,t,g, or c
US-09-487-566-4721
Query Match 86.1%; Score 19.8; DB 18; Length 412;
Best Local Similarity 91.3%; Pred. No. 70;
Matches 21; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CCAACTGGTGGCCATTCAGCTTC 23
Db 253 CCACTGGTGGCCATTCAGCTTC 231

RESULT 9
US-09-487-566A-4721/c
; Sequence 4721, Application US/09487566A
; GENERAL INFORMATION:
; APPLICANT: Rosen, et. al.
; TITLE OF INVENTION: Human Genes, Sequences, and Expression Products 44
; FILE REFERENCE: PO-44
; CURRENT APPLICATION NUMBER: US/09/487,566A
; CURRENT FILING DATE: 2000-01-19
; PRIOR APPLICATION NUMBER: 60/116,668
; PRIOR FILING DATE: 1999-01-20
; NUMBER OF SEQ ID NOS: 5508
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4721
; LENGTH: 412
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (163)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (282)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (313)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (387)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (395)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (402)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (412)
; OTHER INFORMATION: n equals a,t,g, or c

```

US-09-487-566A-4721

```

Query Match 86.1%; Score 19.8; DB 18; Length 412;
Best Local Similarity 91.3%; Pred. No. 70;
Matches 21; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CCAACTGGTGGCCATTCAGCTTC 23
Db 253 CCACTGGTGGCCATTCAGCTTC 231

RESULT 10
US-08-792-586-517/c
; Sequence 517, Application US/08792586
; GENERAL INFORMATION:
; APPLICANT: Stuart, Susan G.
; APPLICANT: Gooding, Douglas H.
; APPLICANT: Lane, John C.
; APPLICANT: Delegeane, Angelo M.
; APPLICANT: Snable, James L.
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM
; TITLE OF INVENTION: COLON WITH CROHN'S DISEASE
; NUMBER OF SEQUENCES: 2231
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/792,586
; FILING DATE: Filed Herewith
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/010,832
; FILING DATE: January 30, 1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/010,812
; FILING DATE: January 30, 1996
; ATTORNEY/AGENT INFORMATION:
; NAME: CERRONE, MICHAEL C.
; REGISTRATION NUMBER: 39,132
; REFERENCE/DOCKET NUMBER: PD-0110 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 855-0555
; TELEFAX: (415) 845-4166
; INFORMATION FOR SEQ ID NO: 517:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 235 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; IMMEDIATE SOURCE:
; CLONE: 775742
; US-08-792-586-517

```

```

Query Match 84.3%; Score 19.4; DB 11; Length 235;
Best Local Similarity 95.2%; Pred. No. 99;
Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 CCAACTGGTGGCCATTCAGCTT 22
Db 224 CCACTGGTGGCCATTCAGCTT 204

```

RESULT 11

US-09-534-843-43302/c
; Sequence 43302, Application US/09534843
; GENERAL INFORMATION:
; APPLICANT: Seilhamer, Jeffrey J.
; APPLICANT: Deleageane, Angelo M.
; APPLICANT: Stuart, Susan G.
; APPLICANT: Stuve, Laura L.
; APPLICANT: Mullahy, Sara J.
; APPLICANT: Naughton, Rebecca E.
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING OR REGULATING SIGNAL TRANSDUCTION MOLECULE
; FILE REFERENCE: PD-1007 CIP
; CURRENT APPLICATION NUMBER: US/09/534,843
; CURRENT FILING DATE: 2000-03-24
; Prior application data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 49783
; SOFTWARE: PERL Program
; SEQ ID NO 43302
; LENGTH: 235
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No: hu00282360
; NAME/KEY: unsure
; LOCATION: 178, 197, 226
; OTHER INFORMATION: a, t, c, g, or other
US-09-534-843-43302

Query Match 84.3%; Score 19.4; DB 20; Length 235;
Best Local Similarity 95.2%; Pred. No. 99;
Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 CAACGTGGTGGCCCTTCAGCTT 22
DB 224 CAACGTGGTGGCCCTTCAGCTT 204

RESULT 12
US-09-107-426-998/c
; Sequence 998, Application US/09107426
; GENERAL INFORMATION:
; APPLICANT: Gooding, Douglas H.
; APPLICANT: Stuart, Susan G.
; APPLICANT: Ito, Laura Y.
; APPLICANT: Naughton, Rebecca E.
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM
; NUMBER OF SEQUENCES: 1626
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/107,426
; FILING DATE: HEREWITH
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: CERRONE, MICHAEL C.
; REGISTRATION NUMBER: 39,132
; REFERENCE/DOCKET NUMBER: PD-0396P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 855-0555
; TELEFAX: (415) 845-4166
; INFORMATION FOR SEQ ID NO: 998:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 235 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; IMMEDIATE SOURCE:
; CLONE: 3421335H1
; US-09-107-426-998

Query Match 81.7%; Score 18.8; DB 49; Length 246;
Best Local Similarity 90.9%; Pred. No. 1.9e+02;
Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CCAACTGGTGGCCCTTCAGCTT 22
DB 204 CCAACTGGTGGCCCTTCAGCTT 183

RESULT 14

; LENGTH: 246 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; IMMEDIATE SOURCE:
; CLONE: 3421335H1
; US-09-107-426-998

Query Match 81.7%; Score 18.8; DB 15; Length 246;
Best Local Similarity 90.9%; Pred. No. 1.9e+02;
Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CCAACTGGTGGCCCTTCAGCTT 22
DB 204 CCAACTGGTGGCCCTTCAGCTT 183

RESULT 13
US-60-051-750-998/c
; Sequence 998, Application US/60051750
; GENERAL INFORMATION:
; APPLICANT: Gooding, Douglas H.
; APPLICANT: Stuart, Susan G.
; APPLICANT: Ito, Laura Y.
; APPLICANT: Naughton, Rebecca E.
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM
; NUMBER OF SEQUENCES: 1626
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/60/051,750
; FILING DATE: HEREWITH
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: CERRONE, MICHAEL C.
; REGISTRATION NUMBER: 39,132
; REFERENCE/DOCKET NUMBER: PD-0396P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 855-0555
; TELEFAX: (415) 845-4166
; INFORMATION FOR SEQ ID NO: 998:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 246 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; IMMEDIATE SOURCE:
; CLONE: 3421335H1
; US-60-051-750-998

Query Match 81.7%; Score 18.8; DB 49; Length 246;
Best Local Similarity 90.9%; Pred. No. 1.9e+02;
Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CCAACTGGTGGCCCTTCAGCTT 22
DB 204 CCAACTGGTGGCCCTTCAGCTT 183

RESULT 14

us-09-270-437d-12.rnrm

Mon Apr 21 10:30:15 2003

US-09-726-807-731/c
; Sequence 731, Application US/09726807
; GENERAL INFORMATION:
; APPLICANT: Gearing, David P.
; APPLICANT: Holtzman, Douglas A.
; APPLICANT: Kingsbury, Gillian A.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID MOLECULES AND USES
; TITLE OF INVENTION: THEREFOR
; FILE REFERENCE: 1600.2053-001
; CURRENT APPLICATION NUMBER: US/09/726,807
; CURRENT FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: 60/168,040
; PRIOR FILING DATE: 1999-11-30
; NUMBER OF SEQ ID NOS: 4076
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 731
; LENGTH: 328
; TYPE: DNA
; ORGANISM: Homo sapiens.
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(328)
; OTHER INFORMATION: n = A,T,C or G
US-09-726-807-731

Query Match 81.7%; Score 18.8; DB 29; Length 328;
Best Local Similarity 90.9%; Pred. No. 2e+02; 2; Indels 0; Gaps 0;
Matches 20; Conservative 0; Mismatches 2

Qy 1 CCAACTGGTGGCCATTGAGCTT 22
Db 145 CCAGCTGGTGGCCCTTCAGCTT 124

RESULT 15
US-08-401-791A-10518/c
; Sequence 10518, Application US/08401791A
; GENERAL INFORMATION:
; APPLICANT: Haseltine, William
; APPLICANT: Rosen, Craig
; APPLICANT: Ruben, Steve
; APPLICANT: Dillon, Patrick
; APPLICANT: Li, Haodong
; APPLICANT: Earle-Hughes, Julie
; TITLE OF INVENTION: Human Genes, Sequences, and Expression Products
; NUMBER OF SEQUENCES: 17621
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Carella, Byrne, Bain, Gilfillan, Cecchi, Stewart, & Olstein
; STREET: 6 Becker Farm Road
; CITY: Roseland
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07068
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 5.0
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/401,791A
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Olstein, Elliot
; REGISTRATION NUMBER: 24,025
; REFERENCE/DOCKET NUMBER: 325800-307
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (201) 994-1700
; TELEFAX: (201) 994-1744
; INFORMATION FOR SEQ ID NO: 10518:

; SEQUENCE CHARACTERISTICS:
; LENGTH: 338 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
US-08-401-791A-10518

Query Match 81.7%; Score 18.8; DB 8; Length 338;
Best Local Similarity 90.9%; Pred. No. 2.1e+02;
Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CCAACTGGTGGCCATTGAGCTT 22
Db 227 CCAGCTGGTGGCCCTTCAGCTT 206

Search completed: April 18, 2003, 09:35:46
Job time : 1268.05 secs

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 18, 2003, 06:35:28 ; Search time 256.311 Seconds
(without alignments)
400.770 Million cell updates/sec

Title: US-09-270-437D-12
Perfect score: 23
Sequence: 1 ccaactgggtggccattcagcttc 23

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 5897297 seqs, 2233080881 residues

Total number of hits satisfying chosen parameters: 11794594

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Pending Patents NA New.*

- 1: /cgn2_6/ptodata/1/pna/PCT_NEW_COMB.seq.*
- 2: /cgn2_6/ptodata/1/pna/US06_NEW_COMB.seq.*
- 3: /cgn2_6/ptodata/1/pna/US07_NEW_COMB.seq.*
- 4: /cgn2_6/ptodata/1/pna/US08_NEW_COMB.seq.*
- 5: /cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq.*
- 6: /cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq2.*
- 7: /cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq3.*
- 8: /cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq2.*
- 9: /cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq3.*
- 10: /cgn2_6/ptodata/1/pna/US60_NEW_COMB.seq.*
- 11: /cgn2_6/ptodata/1/pna/US60_NEW_COMB.seq2.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	23	100.0	23	5	US-09-270-437D-12
2	23	100.0	1708	5	Sequence 12, Appl
3	23	100.0	1946	5	Sequence 5, Appl
4	23	100.0	1990	6	Sequence 7, Appl
5	23	100.0	1990	6	Sequence 7, Appl
6	23	100.0	1990	6	Sequence 7, Appl
7	23	100.0	2040	6	Sequence 6, Appl
8	23	100.0	2040	6	Sequence 6, Appl
9	23	100.0	2069	6	Sequence 6, Appl
10	23	100.0	2069	6	Sequence 10, Appl
11	23	100.0	2119	6	Sequence 10, Appl
12	23	100.0	2158	6	Sequence 5, Appl
13	23	100.0	2158	6	Sequence 5, Appl
14	23	100.0	2208	6	Sequence 9, Appl
15	23	100.0	2208	6	Sequence 9, Appl
16	23	100.0	2427	8	Sequence 8, Appl
17	19.8	86.1	412	6	Sequence 10107, A
18	18.8	81.7	705	6	Sequence 212894, A
19	18.4	80.0	43537	7	Sequence 39653, A
20	17.8	77.4	1396	9	Sequence 13458, A
21	17.8	77.4	3628	1	Sequence 2590, Ap
22	17.8	77.4	3628	9	Sequence 4108, Ap

C 23	17.8	77.4	3628	11	US-60-436-643-4108	Sequence 4108, Ap
C 24	17.2	74.8	201	11	US-60-452-680-41211	Sequence 41211, A
C 25	17.2	74.8	336	8	US-10-401-229-9790	Sequence 9790, Ap
C 26	17.2	74.8	336	8	US-10-401-229-9790	Sequence 21037, A
C 27	17.2	74.8	336	8	US-10-401-229-9790	Sequence 5222, Ap
C 28	17.2	74.8	336	8	US-10-401-229-9790	Sequence 20973, A
C 29	17.2	74.8	429	8	US-10-401-229-9654	Sequence 9654, Ap
C 30	17.2	74.8	537	8	US-10-401-229-9654	Sequence 9654, Ap
C 31	17.2	74.8	601	7	US-09-947-907-9990	Sequence 9990, Ap
C 32	17.2	74.8	601	7	US-09-949-016-33068	Sequence 33068, A
C 33	17.2	74.8	601	7	US-09-949-016-33068	Sequence 33069, A
C 34	17.2	74.8	601	7	US-09-949-016-33070	Sequence 33071, A
C 35	17.2	74.8	601	7	US-09-949-016-33071	Sequence 135882, A
C 36	17.2	74.8	601	7	US-09-949-016-135882	Sequence 135883, A
C 37	17.2	74.8	601	7	US-09-949-016-135883	Sequence 135884, A
C 38	17.2	74.8	601	7	US-09-949-016-135884	Sequence 135885, A
C 39	17.2	74.8	601	7	US-09-949-016-135885	Sequence 752, App
C 40	17.2	74.8	1019	9	US-10-329-624-752	Sequence 1297, Ap
C 41	17.2	74.8	1111	9	US-10-305-720-1297	Sequence 18990, A
C 42	17.2	74.8	1190	8	US-10-170-235-18990	Sequence 13, Appl
C 43	17.2	74.8	1203	1	PCT-IL02-00870-13	Sequence 233, App
C 44	17.2	74.8	1203	6	US-09-949-004-233	Sequence 17364, A
C 45	17.2	74.8	1203	8	US-10-170-235-17364	

ALIGNMENTS

RESULT 1

US-09-270-437D-12
; Sequence 12, Application US/09270437D
; GENERAL INFORMATION:
; APPLICANT: Chen, Yao-Tsang
; APPLICANT: Gure, Ali
; APPLICANT: Tsang, Solam
; APPLICANT: Stockert, Elisabeth
; APPLICANT: Jager, Elke
; APPLICANT: Knuth, Alexander
; APPLICANT: Old, Lloyd J.
; TITLE OF INVENTION: Isolated Nucleic Acid Molecules Encoding Cancer Associated Antigen
; FILE REFERENCE: LUD 5538.1
; CURRENT APPLICATION NUMBER: US/09/270,437D
; PRIOR FILING DATE: 1999-03-16
; PRIOR APPLICATION NUMBER: 09/061,709
; NUMBER OF SEQ ID NOS: 23
; SEQ ID NO 12
; LENGTH: 23
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
US-09-270-437D-12

Query Match 100.0%; Score 23; DB 5; Length 23;
Best Local Similarity 100.0%; Pred. No. 0.34; Indels 0; Gaps 0;
Matches 23; Conservative 0; Mismatches 0

Qy 1 CCAACTGGTGGCCATTGAGCTTC 23

Db 1 CCAACTGGTGGCCATTGAGCTTC 23

RESULT 2

US-09-270-437D-5/c
; Sequence 5, Application US/09270437D
; GENERAL INFORMATION:
; APPLICANT: Chen, Yao-Tsang
; APPLICANT: Gure, Ali
; APPLICANT: Tsang, Solam
; APPLICANT: Stockert, Elisabeth
; APPLICANT: Jager, Elke
; APPLICANT: Knuth, Alexander

us-09-270-437d-12.rnpn

Mon Apr 21 10:30:17 2003

```

; APPLICANT: Old, Lloyd J.
; TITLE OF INVENTION: Isolated Nucleic Acid Molecules Encoding Cancer Associated Antigen
; FILE REFERENCE: LUD 538.1
; CURRENT APPLICATION NUMBER: US/09/270,437D
; PRIOR FILING DATE: 1999-03-16
; PRIOR APPLICATION NUMBER: 09/061,709
; PRIOR FILING DATE: 1998-04-17
; NUMBER OF SEQ ID NOS: 23
; SEQ ID NO 5
; LENGTH: 1708
; TYPE: DNA
; ORGANISM: Homo sapiens
; NAME/KEY: CDS
; LOCATION: 1384,1464,1533,1571,1595
; OTHER INFORMATION: unsure of nucleotide
US-09-270-437D-5

```

```

Query Match      100.0%; Score 23; DB 5; Length 1708;
Best Local Similarity 100.0%; Pred. No. 0.54;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy 1 CCAACTGGTGGCCATTGAGCTTC 23
    |||||
Db 68 CCAACTGGTGGCCATTGAGCTTC 46

```

```

RESULT 3
US-09-270-437D-7/c
; Sequence 7, Application US/09270437D
; GENERAL INFORMATION:
; APPLICANT: Chen, Yao-Tseung
; APPLICANT: Gure, Ali
; APPLICANT: Tsang, Solam
; APPLICANT: Stockert, Elisabeth
; APPLICANT: Jager, Elke
; APPLICANT: Knuth, Alexander
; APPLICANT: Old, Lloyd J.
; TITLE OF INVENTION: Isolated Nucleic Acid Molecules Encoding Cancer Associated Antigen
; FILE REFERENCE: LUD 538.1
; CURRENT APPLICATION NUMBER: US/09/270,437D
; PRIOR FILING DATE: 1999-03-16
; PRIOR APPLICATION NUMBER: 09/061,709
; PRIOR FILING DATE: 1998-04-17
; NUMBER OF SEQ ID NOS: 23
; SEQ ID NO 7
; LENGTH: 1946
; TYPE: DNA
; ORGANISM: Homo sapiens
; NAME/KEY: CDS
; LOCATION: 1622,1702,1771,1809,1833
; OTHER INFORMATION: unsure of nucleotide
US-09-270-437D-7

```

```

Query Match      100.0%; Score 23; DB 5; Length 1946;
Best Local Similarity 100.0%; Pred. No. 0.55;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy 1 CCAACTGGTGGCCATTGAGCTTC 23
    |||||
Db 306 CCAACTGGTGGCCATTGAGCTTC 284

```

```

RESULT 4
US-09-724-676-7/c
; Sequence 7, Application US/09724676
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen

```

```

; CURRENT APPLICATION NUMBER: US/09/724,676
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 7
; LENGTH: 1990
; TYPE: DNA
; ORGANISM: Homo sapiens
; OTHER INFORMATION: PatentIn version 3.2
US-09-724-676-7

```

```

Query Match      100.0%; Score 23; DB 6; Length 1990;
Best Local Similarity 100.0%; Pred. No. 0.55;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy 1 CCAACTGGTGGCCATTGAGCTTC 23
    |||||
Db 442 CCAACTGGTGGCCATTGAGCTTC 420

```

```

RESULT 5
US-09-724-676A-7/c
; Sequence 7, Application US/09724676A
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676A
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 7
; LENGTH: 1990
; TYPE: DNA
; ORGANISM: Homo sapiens
; OTHER INFORMATION: PatentIn version 3.2
US-09-724-676A-7

```

```

Query Match      100.0%; Score 23; DB 6; Length 1990;
Best Local Similarity 100.0%; Pred. No. 0.55;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy 1 CCAACTGGTGGCCATTGAGCTTC 23
    |||||
Db 442 CCAACTGGTGGCCATTGAGCTTC 420

```

```

RESULT 6
US-09-724-676-6/c
; Sequence 6, Application US/09724676
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 6
; LENGTH: 2040
; TYPE: DNA
; ORGANISM: Homo sapiens
; OTHER INFORMATION: PatentIn version 3.2
US-09-724-676-6

```

```

Query Match      100.0%; Score 23; DB 6; Length 2040;
Best Local Similarity 100.0%; Pred. No. 0.55;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy 1 CCAACTGGTGGCCATTGAGCTTC 23
    |||||
Db 442 CCAACTGGTGGCCATTGAGCTTC 420

```

```

RESULT 7
US-09-724-676A-6/c

```



```
; Sequence 6, Application US/09724676A
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676A
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 6
; LENGTH: 2040
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-724-676A-6

Query Match      100.0%; Score 23; DB 6; Length 2040;
Best Local Similarity 100.0%; Pred. No. 0.55;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCAACTGGTGGCCATTGAGCTTC 23
|||||
DB 442 CCAACTGGTGGCCATTGAGCTTC 420

RESULT 8
US-09-724-676-10/c
; Sequence 10, Application US/09724676
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 10
; LENGTH: 2069
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-724-676-10

Query Match      100.0%; Score 23; DB 6; Length 2069;
Best Local Similarity 100.0%; Pred. No. 0.55;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCAACTGGTGGCCATTGAGCTTC 23
|||||
DB 442 CCAACTGGTGGCCATTGAGCTTC 420

RESULT 9
US-09-724-676A-10/c
; Sequence 10, Application US/09724676A
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676A
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 10
; LENGTH: 2069
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-724-676A-10

Query Match      100.0%; Score 23; DB 6; Length 2069;
Best Local Similarity 100.0%; Pred. No. 0.55;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCAACTGGTGGCCATTGAGCTTC 23
|||||
DB 442 CCAACTGGTGGCCATTGAGCTTC 420

RESULT 10
US-09-724-676-5/c
; Sequence 5, Application US/09724676
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 5
; LENGTH: 2119
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-724-676-5

Query Match      100.0%; Score 23; DB 6; Length 2119;
Best Local Similarity 100.0%; Pred. No. 0.55;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCAACTGGTGGCCATTGAGCTTC 23
|||||
DB 442 CCAACTGGTGGCCATTGAGCTTC 420

RESULT 11
US-09-724-676A-5/c
; Sequence 5, Application US/09724676A
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676A
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 5
; LENGTH: 2119
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-724-676A-5

Query Match      100.0%; Score 23; DB 6; Length 2119;
Best Local Similarity 100.0%; Pred. No. 0.55;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCAACTGGTGGCCATTGAGCTTC 23
|||||
DB 442 CCAACTGGTGGCCATTGAGCTTC 420

RESULT 12
US-09-724-676-9/c
; Sequence 9, Application US/09724676
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 9
; LENGTH: 2158
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-724-676-9

Query Match      100.0%; Score 23; DB 6; Length 2158;
```

	Best Local Similarity	100.0%;	Pred. No. 0.56;	Indels	0;	Gaps	0;
	Matches	23;	Conservative	Mismatches	0;		
QY	1	CCAACTGGTGGCCATT	CAGCTTC	23			
Db	442	CCAACTGGTGGCCATT	CAGCTTC	420			
RESULT 13							

GENERAL INFORMATION: LTD
APPLICANT: Compugen
TITLE OR INVENTION: Variants of alternative splicing

```

; FILE REFERENCE: 12301014 COMPACT
; CURRENT APPLICATION NUMBER: US/09/724,676A
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2

```

```
; SEQ ID NO 9
; LENGTH: 2158
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-724-676A-9
```

Query Match	100.0%;	Score 23;	DB 6;	Length 2158;
Best Local Similarity	100.0%;	Pred. No. 0.56;		
			Indels	0; Gaps 0;

QY	1	C C A A C T G G T G G C C A T T C A G C T T C	23
D5-	442	C C A A C T G G T G G C C A T T C A G C T T C	420

```

RESULT 14
US-09-724-676-8/c
; Sequence 8, Application US/09724676
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 8
; LENGTH: 2208
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-724-676-8

```

Query Match	100.0%	Score 23	DB 6	Length 2208
Best Local Similarity	100.0%	Pred. NO.	0.56	
Basech 23	Conservative	0	Mismatches	0
		0	Indels	0
		0	Gaps	0

[illegible]

```

RESULT 15
US-09-724-676A-8/c
; Sequence 8 Application US/09724676A
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676A
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 8
; LENGTH: 2208

```

GenCore version 5.1.4.p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 18, 2003, 04:47:40 ; Search time 800.818 Seconds
(without alignments)
465.145 Million cell updates/sec

Title: .US-09-270-437D-12

Perfect score: 23

Sequence: 1 ccaactgggtggccattcagcttc 23

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

1: em_estba:*

2: em_esthum:*

3: em_estin:*

4: em_estmu:*

5: em_estov:*

6: em_estpl:*

7: em_estro:*

8: em_hic:*

9: gb_est1:*

10: gb_est2:*

11: gb_hic:*

12: gb_est3:*

13: gb_est4:*

14: gb_est5:*

15: em_estfun:*

16: em_estom:*

17: gb_gss:*

18: em_gss_hum:*

19: em_gss_inv:*

20: em_gss_pln:*

21: em_gss_vrt:*

22: em_gss_fun:*

23: em_gss_man:*

24: em_gss_mus:*

25: em_gss_other:*

26: em_gss_pro:*

27: em_gss_rod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	23	100.0	249	9	AA196774
C 2	23	100.0	466	9	AA196976
C 3	23	100.0	764	13	BM006191
C 4	23	100.0	877	12	BG473949
C 5	23	100.0	896	14	BQ647425
C 6	23	100.0	929	14	BQ652586

C 7	23	100.0	940	14	BQ649666
C 8	23	100.0	950	14	BQ648383
C 9	23	100.0	953	14	BQ647360
C 10	23	100.0	955	14	BQ651665
C 11	23	100.0	982	14	BQ652180
C 12	23	100.0	983	14	BQ643920
C 13	23	100.0	1042	14	BM928196
C 14	23	100.0	1061	14	BQ647561
C 15	21.4	93.0	770	10	BES41467
C 16	19.4	84.3	912	17	B08385
C 17	18.8	81.7	326	13	BM148903
C 18	18.8	81.7	332	17	AQ132664
C 19	18.8	81.7	505	13	BM146309
C 20	18.8	81.7	589	12	BG025694
C 21	18.8	81.7	641	12	B797894
C 22	18.8	81.7	844	17	AZ901563
C 23	18.8	81.7	1095	13	BM475738
C 24	18.2	79.1	155	17	BH315184
C 25	18.2	79.1	433	10	BES08123
C 26	18.2	79.1	508	12	BG407307
C 27	18.2	79.1	528	13	BI348247
C 28	18.2	79.1	588	13	BI446082
C 29	18.2	79.1	719	17	BH444983
C 30	18.2	79.1	776	17	BH054219
C 31	17.8	77.4	325	10	AV952715
C 32	17.8	77.4	414	17	AZ762364
C 33	17.8	77.4	482	13	BI401056
C 34	17.8	77.4	551	17	B26736
C 35	17.8	77.4	552	10	BZ231931
C 36	17.8	77.4	570	14	BQ603654
C 37	17.8	77.4	604	10	AV825598
C 38	17.8	77.4	621	9	AL683579
C 39	17.8	77.4	634	9	AL595504
C 40	17.8	77.4	657	9	AU237120
C 41	17.8	77.4	682	9	AL640616
C 42	17.8	77.4	724	12	BF307752
C 43	17.8	77.4	796	17	BI1822
C 44	17.4	75.7	401	17	AZ000183
C 45	17.4	75.7	597	17	AQ482316

ALIGNMENTS

RESULT 1
AA196774/c
LOCUS
DEFINITION
zq60f02.r1 Stratagene neuroepithelium (#937231) Homo sapiens CDNA
clone IMAGE:646011 5', mRNA sequence.
ACCESSION
AA196774
VERSION
AA196774.1 GI:1792476
KEYWORDS
EST.
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 249)
Hillier, L., Lennon, G., Becker, M., Bonaldo, M.F., Chiapelli, B.,
Chisoe, S., Dietrich, N., Dubuque, T., Favello, A., Gish, W., Hawkins,
M., Hultman, M., Kucaba, T., Lacy, M., Le, M., Le, N., Mardis, E., Moore,
B., Morris, M., Parsons, J., Prange, C., Rifkin, L., Rohlfing, T.,
Schellenberg, K., Soares, M.B., Tan, F., Thierry-Mieg, J., Trevaskis, E.,
Underwood, K., Wohlmann, P., Waterston, R., Wilson, R. and Marra, M.,
Generation and analysis of 280,000 human expressed sequence tags
Genome Res. 6 (9), 807-828 (1996)
Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available royalty-free through LLNL ; contact the

BQ649666 AGENCOURT
BQ648383 AGENCOURT
BQ647360 AGENCOURT
BQ651665 AGENCOURT
BQ652180 AGENCOURT
BQ643920 AGENCOURT
BM928196 AGENCOURT
BQ647561 AGENCOURT
BES41467 601067919
B08385 F18A15-T7.1
BM148903 TCAAP2852
AQ132664 HS 3058_A
BM146309 TCAAP1543
BG025694 602274861
BF797894 602257293
AZ901563 RPCI-24-2
BM475738 AGENCOURT
BH315184 CH230-209
BES08123 dc26h02.y
BG407307 dabi7d05.y
BI348247 dai127a07.
BI446082 dai135908.
BH444983 BQB846TR
BH054219 RPCI-24-3
AV952715 AV952715
AZ762364 LM0557F21
BI401056 MI-P-CP0-
B26736 T18CTTF TAM
BZ231931 136808 MA
BQ603654 MI-P-CP0-
AV825598 AV825598
AL683579 AL683579
AL595504 AL595504
AU237120 AU237120
AL640616 AL640616
BF307752 601890820
BI1822 F18A15-T7.3
AZ000183 RPCI-23-3
AQ482316 RPCI-11-2

us-09-270-437d-12.rst

Mon Apr 21 10:30:18 2003

cells (Ntera-2/ci.D1) induced with Retinoic Acid for 24 hours. Average insert size: 1.5 kb; Uni-ZAP XR Vector; ~5' adaptor sequence: 5' GAATTCGGCAGAG 3' ~3' adaptor sequence: 5' CTCGAGTTTTTTTTTTT 3' others

BASE COUNT 121 a 123 c 124 g 87 t 11 others

FEATURES source

1. .249
/organism="Homo sapiens"
/db_xref="GDB:5216262"
/db_xref="taxon:9606"
/clone="IMAGE:646011"
/clone_lib="Stratagene neuroepithelium (#937231)"
/dev_stage="Ntera-2/RA neuroepithelial cells"
/lab_host="SOLR (kanamycin resistant)"
/note="vector: pBluescript SK-; Site 1: ECORI; Site 2: XhoI; Cloned unidirectionally. Primer: Oligo dt. NT2 cells (Ntera-2/ci.D1) induced with Retinoic Acid for 24 hours. Average insert size: 1.5 kb; Uni-ZAP XR Vector; ~5' adaptor sequence: 5' GAATTCGGCAGAG 3' ~3' adaptor sequence: 5' CTCGAGTTTTTTTTTTT 3' others

BASE COUNT 56 a 71 c 73 g 43 t 6 others

Query Match 100.0%; Score 23; DB 9; Length 249;
Best Local Similarity 100.0%; Pred. No. 2.8; 0; Indels 0; Gaps 0;
Matches 23; Conservative 0; Mismatches 0;

Qy 1 CCAACTGGTGGCCATTCAGCTTC 23
Db 42 CCAACTGGTGGCCATTCAGCTTC 20

RESULT 2
AA196976/c 456 bp mRNA linear EST 22-JAN-1997
LOCUS zq60h02.r1 Stratagene neuroepithelium (#937231) Homo sapiens CDNA
DEFINITION clone IMAGE:646035 5', mRNA sequence.

ACCESSION AA196976
VERSION AA196976.1 GI:1792567
KEYWORDS EST.
SOURCE human.

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 466)
Hillier, L., Lennon, N., Dubuque, T., Pavello, A., Gish, W., Hawkins, Chissoe, S., Dietrich, G., Kucaba, T., Lacy, M., Le, M., Mardis, E., Moore, B., Hultman, M., Parsons, J., Prange, C., Rifkin, L., Rohlfing, T., Schellenberg, K., Soares, M.B., Tan, F., Thierry-Weg, J., Trevaaskis, E., Underwood, K., Wohldmann, P., Waterston, R., Wilson, R. and Marra, M. Generation and analysis of 280,000 human expressed sequence tags
Genome Res. 6 (9), 807-828 (1996)

TITLE 97044478
JOURNAL Contact: Wilson RK
MEDLINE Washington University School of Medicine
COMMENT 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810

Email: est@watson.wustl.edu
This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Seq primer: -28M13 rev2 from Amersham
High quality sequence stop: 167.

FEATURES source

1. .466
/organism="Homo sapiens"
/db_xref="GDB:5216262"
/db_xref="taxon:9606"
/clone="IMAGE:646035"
/clone_lib="Stratagene neuroepithelium (#937231)"
/dev_stage="Ntera-2/RA neuroepithelial cells"
/lab_host="SOLR (kanamycin resistant)"
/note="vector: pBluescript SK-; Site 1: ECORI; Site 2: XhoI; Cloned unidirectionally. Primer: Oligo dt. NT2

Query Match 100.0%; Score 23; DB 9; Length 466;
Best Local Similarity 100.0%; Pred. No. 3.7; 0; Indels 0; Gaps 0;
Matches 23; Conservative 0; Mismatches 0;

Qy 1 CCAACTGGTGGCCATTCAGCTTC 23
Db 42 CCAACTGGTGGCCATTCAGCTTC 20

RESULT 3
BM006191/c 764 bp mRNA linear EST 30-OCT-2001
LOCUS 603614013F1 NIH_MGC_100 Homo sapiens CDNA clone IMAGE:5433698 5',
DEFINITION mRNA sequence.

ACCESSION BM006191
VERSION BM006191.1 GI:16520545
KEYWORDS EST.
SOURCE human.

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 764)
NIH-MGC http://mgc.ncbi.nlm.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov

Tissue procurement: CGAP (Stanford)
CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCM1909 row: j column: 03
High quality sequence stop: 756.

FEATURES source

Location/Qualifiers
1. .764
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5433698"
/clone_lib="NIH_MGC_100"
/tissue_type="hepatocellular carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: liver; Vector: pOTB7; Site 1: XhoI; Site 2: ECORI; CDNA made by oligo-dt priming. Directionally cloned into ECORI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Size-selected 500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin. (University of California, Berkeley) using ZAP-CDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library."
198 a 209 c 217 g 140 t

BASE COUNT 198 a 209 c 217 g 140 t

Query Match 100.0%; Score 23; DB 13; Length 764;
Best Local Similarity 100.0%; Pred. No. 4.6; 0; Indels 0; Gaps 0;
Matches 23; Conservative 0; Mismatches 0;

Qy 1 CCAACTGGTGGCCATTCAGCTTC 23
Db 45 CCAACTGGTGGCCATTCAGCTTC 23

RESULT 4

BG473949/c
 LOCUS BG473949 877 bp mRNA linear EST 21-MAR-2001
 DEFINITION 60251644F1 NIH_MGC_16 Homo sapiens cDNA clone IMAGE:4647994 5',
 mRNA sequence.
 ACCESSION BG473949
 VERSION BG473949.1 GI:13406226
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 877)
 AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-r@mail.nih.gov
 Tissue Procurement: ATCC
 cDNA Library Preparation: Ling Hong/Rubin Laboratory
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Cloned into EcoRI/XhoI sites using the following 5' adaptor:
 GGACGAG(G). Library constructed by Ling Hong in the laboratory
 in the laboratory of Gerald M. Rubin (University of California, Berkeley)
 using ZAP-cDNA synthesis kit (Stratagene) and Superscript
 II RT (Life Technologies). Note: this is a NIH_MGC Library.
 found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
 Plate: LLCM1425 row: h column: 11
 High quality sequence stop: 673.
 Location/Qualifiers
 1..877
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:4647994"
 /clone_lib="NIH_MGC_16"
 /tissue_type="retinoblastoma"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: eye; Vector: pOTB7; Site 1: XhoI; Site 2:
 EcoRI; cDNA made by oligo-dt priming. Directionally
 cloned into EcoRI/XhoI sites using the following 5'
 adaptor: GGACGAG(G). Library constructed by Ling Hong
 in the laboratory of Gerald M. Rubin (University of
 California, Berkeley) using ZAP-cDNA synthesis kit
 (Stratagene) and Superscript II RT (Life Technologies).
 Note: this is a NIH_MGC Library."
 BASE COUNT 205 a 269 c 257 g 146 t
 ORIGIN
 Query Match 100.0%; Score 23; DB 12; Length 877;
 Best Local Similarity 100.0%; Pred. No. 4.9;
 Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 CCAACTGGTGGCCATTCAGCTTC 23
 Db 690 CCAACTGGTGGCCATTCAGCTTC 668
 RESULT 5
 BG647425/c
 LOCUS BG647425 896 bp mRNA linear EST 15-JUL-2002
 DEFINITION AGENCOURT_8346875 NIH_MGC_100 Homo sapiens cDNA clone IMAGE:6283876
 5', mRNA sequence.
 ACCESSION BG647425
 VERSION BG647425.1 GI:21771597
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 896)
 AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-r@mail.nih.gov
 Tissue Procurement: CGAP (Stanford)
 cDNA Library Preparation: Rubin Laboratory
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Cloned into EcoRI/XhoI sites using the following 5' adaptor:
 GGACGAG(G). Library constructed by Ling Hong in the laboratory
 in the laboratory of Gerald M. Rubin (University of California, Berkeley)
 using ZAP-cDNA synthesis kit (Stratagene) and Superscript
 II RT (Life Technologies). Note: this is a NIH_MGC Library.
 found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 Plate: LLCM2497 row: k column: 24
 High quality sequence stop: 533.
 Location/Qualifiers
 1..929
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:6293903"
 /clone_lib="NIH_MGC_100"
 /tissue_type="hepatocellular carcinoma, cell line"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: liver; Vector: pOTB7; Site 1: XhoI; Site 2:
 EcoRI; cDNA made by oligo-dt priming. Directionally cloned

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Cloned into EcoRI/XhoI sites using the following 5' adaptor:
 GGACGAG(G). Size-selected >500bp for average insert size
 1.8kb. Library constructed by Ling Hong in the laboratory
 of Gerald M. Rubin (University of California, Berkeley)
 using ZAP-cDNA synthesis kit (Stratagene) and Superscript
 II RT (Life Technologies). Note: this is a NIH_MGC
 Library."
 found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 Plate: LLCM2480 row: j column: 05
 High quality sequence stop: 421.
 Location/Qualifiers
 1..896
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:6283876"
 /clone_lib="NIH_MGC_100"
 /tissue_type="hepatocellular carcinoma, cell line"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: liver; Vector: pOTB7; Site 1: XhoI; Site 2:
 EcoRI; cDNA made by oligo-dt priming. Directionally cloned
 into EcoRI/XhoI sites using the following 5' adaptor:
 GGACGAG(G). Size-selected >500bp for average insert size
 1.8kb. Library constructed by Ling Hong in the laboratory
 of Gerald M. Rubin (University of California, Berkeley)
 using ZAP-cDNA synthesis kit (Stratagene) and Superscript
 II RT (Life Technologies). Note: this is a NIH_MGC
 Library."
 BASE COUNT 255 a 235 c 252 g 150 t 4 others
 ORIGIN
 Query Match 100.0%; Score 23; DB 14; Length 896;
 Best Local Similarity 100.0%; Pred. No. 5;
 Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 CCAACTGGTGGCCATTCAGCTTC 23
 Db 354 CCAACTGGTGGCCATTCAGCTTC 332
 RESULT 6
 BG652586/c
 LOCUS BG652586 929 bp mRNA linear EST 15-JUL-2002
 DEFINITION AGENCOURT_8490818 NIH_MGC_100 Homo sapiens cDNA clone IMAGE:6293903
 5', mRNA sequence.
 ACCESSION BG652586
 VERSION BG652586.1 GI:21776758
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 929)
 AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-r@mail.nih.gov
 Tissue Procurement: CGAP (Stanford)
 cDNA Library Preparation: Rubin Laboratory
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Cloned into EcoRI/XhoI sites using the following 5' adaptor:
 GGACGAG(G). Library constructed by Ling Hong in the laboratory
 in the laboratory of Gerald M. Rubin (University of California, Berkeley)
 using ZAP-cDNA synthesis kit (Stratagene) and Superscript
 II RT (Life Technologies). Note: this is a NIH_MGC Library.
 found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 Plate: LLCM2497 row: k column: 24
 High quality sequence stop: 533.
 Location/Qualifiers
 1..929
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:6293903"
 /clone_lib="NIH_MGC_100"
 /tissue_type="hepatocellular carcinoma, cell line"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: liver; Vector: pOTB7; Site 1: XhoI; Site 2:
 EcoRI; cDNA made by oligo-dt priming. Directionally cloned

```

into EcoRI/XhoI sites using the following 5' adaptor:
GGCAGGAG(G). Size-selected >500bp for average insert size
1.8kb. Library constructed by Ling Hong in the laboratory
of Gerald M. Rubin (University of California, Berkeley)
using ZAP-cDNA synthesis kit (Stratagene) and Superscript
II RT (Life Technologies). Note: this is a NIH_MGC
Library."
BASE COUNT      260 a      247 c      264 g      158 t
ORIGIN
Query Match      100.0%; Score 23; DB 14; Length 929;
Best Local Similarity 100.0%; Pred. No. 5;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CCAACTGGTGGCCATTGAGCTTC 23
|||||
Db 417 CCAACTGGTGGCCATTGAGCTTC 395

RESULT 7
BQ649666/c
LOCUS      BQ649666      940 bp      mRNA      linear      EST 15-JUL-2002
DEFINITION AGENCOURT 8286221 NIH_MGC_100 Homo sapiens cDNA clone IMAGE:6299693
5', mRNA sequence.
ACCESSION  BQ649666
VERSION     BQ649666
KEYWORDS    EST.
SOURCE      human.
ORGANISM    Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 940)
AUTHORS     NIH-MGC http://mgc.nci.nih.gov/.
TITLE       NIH-MGC
JOURNAL
COMMENT     National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: CGAP (Stanford)
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
DNA Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCM2512 row: m column: 06
High quality sequence stop: 395.
FEATURES
Location/Qualifiers
1..940
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:6299693"
/clone_lib="NIH_MGC_100"
/tissue_type="hepatocellular carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/notes="Organ: liver; Vector: pOTB7; Site 1: XhoI; Site 2:
EcoRI; cDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCAGGAG(G). Size-selected >500bp for average insert size
1.8kb. Library constructed by Ling Hong in the laboratory
of Gerald M. Rubin (University of California, Berkeley)
using ZAP-cDNA synthesis kit (Stratagene) and Superscript
II RT (Life Technologies). Note: this is a NIH_MGC
Library."
BASE COUNT      258 a      269 c      262 g      149 t      2 others
ORIGIN
Query Match      100.0%; Score 23; DB 14; Length 940;
Best Local Similarity 100.0%; Pred. No. 5.1;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CCAACTGGTGGCCATTGAGCTTC 23
|||||
Db 337 CCAACTGGTGGCCATTGAGCTTC 315

```

```

RESULT 8
BQ648383/c
LOCUS      BQ648383      950 bp      mRNA      linear      EST 15-JUL-2002
DEFINITION AGENCOURT 8342265 NIH_MGC_100 Homo sapiens cDNA clone IMAGE:6268202
5', mRNA sequence.
ACCESSION  BQ648383
VERSION     BQ648383
KEYWORDS    EST.
SOURCE      human.
ORGANISM    Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 950)
AUTHORS     NIH-MGC http://mgc.nci.nih.gov/.
TITLE       NIH-MGC
JOURNAL
COMMENT     National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: CGAP (Stanford)
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
DNA Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCM2439 row: m column: 03
High quality sequence stop: 599.
FEATURES
Location/Qualifiers
1..950
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:6268202"
/clone_lib="NIH_MGC_100"
/tissue_type="hepatocellular carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/notes="Organ: liver; Vector: pOTB7; Site 1: XhoI; Site 2:
EcoRI; cDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCAGGAG(G). Size-selected >500bp for average insert size
1.8kb. Library constructed by Ling Hong in the laboratory
of Gerald M. Rubin (University of California, Berkeley)
using ZAP-cDNA synthesis kit (Stratagene) and Superscript
II RT (Life Technologies). Note: this is a NIH_MGC
Library."
BASE COUNT      237 a      278 c      269 g      166 t
ORIGIN
Query Match      100.0%; Score 23; DB 14; Length 950;
Best Local Similarity 100.0%; Pred. No. 5.1;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CCAACTGGTGGCCATTGAGCTTC 23
|||||
Db 43 CCAACTGGTGGCCATTGAGCTTC 21

RESULT 9
BQ647360/c
LOCUS      BQ647360      953 bp      mRNA      linear      EST 15-JUL-2002
DEFINITION AGENCOURT 8443628 NIH_MGC_100 Homo sapiens cDNA clone IMAGE:6284231
5', mRNA sequence.
ACCESSION  BQ647360
VERSION     BQ647360
KEYWORDS    EST.
SOURCE      human.
ORGANISM    Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 953)
AUTHORS     NIH-MGC http://mgc.nci.nih.gov/.
TITLE       NIH-MGC

```

JOURNAL
COMMENT

Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-rc@mail.nih.gov
Tissue Procurement: CGAP (Stanford)
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCM2481 row: h column: 24
High quality sequence stop: 651.
Location/Qualifiers
1. .953
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="IMAGE:6284231"
/clone_lib="NIH_MGC_100"
/tissue_type="hepatocellular carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: liver; Vector: pOTB7; Site 1: XhoI; Site 2:
into EcoRI/XhoI sites using the following 5' adaptor:
GGCAGCAG(G). Size-selected >500bp for average insert size
1.8kb. Library constructed by Ling Hong in the laboratory
of Gerald M. Rubin (University of California, Berkeley)
using ZAP-cDNA synthesis kit (Stratagene) and Superscript
II RT (Life Technologies). Note: this is a NIH_MGC
Library."
BASE COUNT 277 a 251 c 258 g 166 t 1 others
LOCUS BQ651665
DEFINITION AGENCOURT 8297761 NIH_MGC_100 Homo sapiens cDNA clone IMAGE:6269842
5', mRNA sequence.
ACCESSION BQ651665
VERSION BQ651665.1 GI:21775837
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NIH-MGC http://mgi.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-rc@mail.nih.gov
Tissue Procurement: CGAP (Stanford)
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCM2444 row: a column: 11
High quality sequence stop: 588.
Location/Qualifiers
1. .955
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="IMAGE:6269842"

FEATURES
source

Query Match 100.0%; Score 23; DB 14; Length 953;
Best Local Similarity 100.0%; Pred. No. 5.1;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 CCAACTGGTGGCCATTTCAGCTTC 23
|||||
Db 336 CCAACTGGTGGCCATTTCAGCTTC 314
|||||
RESULT 10
BQ651665/c
LOCUS BQ651665
DEFINITION AGENCOURT 8297761 NIH_MGC_100 Homo sapiens cDNA clone IMAGE:6269842
5', mRNA sequence.
ACCESSION BQ651665
VERSION BQ651665.1 GI:21775837
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NIH-MGC http://mgi.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-rc@mail.nih.gov
Tissue Procurement: CGAP (Stanford)
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCM2444 row: a column: 11
High quality sequence stop: 588.
Location/Qualifiers
1. .955
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="IMAGE:6269842"

BASE COUNT
ORIGIN

Query Match 100.0%; Score 23; DB 14; Length 953;
Best Local Similarity 100.0%; Pred. No. 5.1;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 CCAACTGGTGGCCATTTCAGCTTC 23
|||||
Db 336 CCAACTGGTGGCCATTTCAGCTTC 314
|||||

RESULT 10
BQ651665/c

LOCUS BQ651665
DEFINITION AGENCOURT 8297761 NIH_MGC_100 Homo sapiens cDNA clone IMAGE:6269842
5', mRNA sequence.
ACCESSION BQ651665
VERSION BQ651665.1 GI:21775837
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NIH-MGC http://mgi.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-rc@mail.nih.gov
Tissue Procurement: CGAP (Stanford)
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCM2444 row: a column: 11
High quality sequence stop: 588.
Location/Qualifiers
1. .955
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="IMAGE:6269842"

REFERENCE
AUTHORS

TITLE NIH-MGC http://mgi.nci.nih.gov/
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-rc@mail.nih.gov
Tissue Procurement: CGAP (Stanford)
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCM2444 row: a column: 11
High quality sequence stop: 588.
Location/Qualifiers
1. .955
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="IMAGE:6269842"

FEATURES
source

Query Match 100.0%; Score 23; DB 14; Length 982;
Best Local Similarity 100.0%; Pred. No. 5.2;
BASE COUNT 292 a 251 c 267 g 170 t 2 others
ORIGIN

/clone_lib="NIH_MGC_100"
/tissue_type="hepatocellular carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: liver; Vector: pOTB7; Site 1: XhoI; Site 2:
into EcoRI/XhoI sites using the following 5' adaptor:
GGCAGCAG(G). Size-selected >500bp for average insert size
1.8kb. Library constructed by Ling Hong in the laboratory
of Gerald M. Rubin (University of California, Berkeley)
using ZAP-cDNA synthesis kit (Stratagene) and Superscript
II RT (Life Technologies). Note: this is a NIH_MGC
Library."
BASE COUNT 232 a 313 c 253 g 157 t
ORIGIN

Query Match 100.0%; Score 23; DB 14; Length 955;
Best Local Similarity 100.0%; Pred. No. 5.1;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 CCAACTGGTGGCCATTTCAGCTTC 23
|||||
Db 43 CCAACTGGTGGCCATTTCAGCTTC 21
|||||

RESULT 11
BQ652180/c

LOCUS BQ652180
DEFINITION AGENCOURT 8188778 NIH_MGC_100 Homo sapiens cDNA clone IMAGE:6282804
5', mRNA sequence.
ACCESSION BQ652180
VERSION BQ652180.1 GI:21776352
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NIH-MGC http://mgi.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-rc@mail.nih.gov
Tissue Procurement: CGAP (Stanford)
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCM2477 row: m column: 13
High quality sequence stop: 667.
Location/Qualifiers
1. .982
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="IMAGE:6282804"
/clone_lib="NIH_MGC_100"
/tissue_type="hepatocellular carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: liver; Vector: pOTB7; Site 1: XhoI; Site 2:
into EcoRI/XhoI sites using the following 5' adaptor:
GGCAGCAG(G). Size-selected >500bp for average insert size
1.8kb. Library constructed by Ling Hong in the laboratory
of Gerald M. Rubin (University of California, Berkeley)
using ZAP-cDNA synthesis kit (Stratagene) and Superscript
II RT (Life Technologies). Note: this is a NIH_MGC
Library."
BASE COUNT 292 a 251 c 267 g 170 t 2 others
ORIGIN

FEATURES
source

Query Match 100.0%; Score 23; DB 14; Length 982;
Best Local Similarity 100.0%; Pred. No. 5.2;
BASE COUNT 292 a 251 c 267 g 170 t 2 others
ORIGIN

us-09-270-437d-12.rst

Mon Apr 21 10:30:18 2003

Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCAACTGGTGGCCATTGAGCTTC 23
 Db 336 CCAACTGGTGGCCATTGAGCTTC 314

RESULT 12
 BQ643920/c 983 bp mRNA linear EST 15-JUL-2002
 LOCUS AGENCOURT 8303966 NIH_MGC_100 Homo sapiens cDNA clone IMAGE:6269413
 DEFINITION 5', mRNA sequence.

ACCESSION BQ643920
 VERSION BQ643920.1 GI:21768092
 KEYWORDS EST.
 SOURCE human.

ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 983)
 NIH-MGC http://mgs.nci.nih.gov/
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: CGAP (Stanford)
 cDNA Library Preparation: Rubin Laboratory
 DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: LLCM2442 row: 0 column: 14
 High quality sequence stop: 491.

FEATURES
 source
 1..983
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:6269413"
 /clone_lib="NIH_MGC_100"
 /tissue_type="hepatocellular carcinoma, cell line"
 /lab_host="PH108 (phage-resistant)"
 /note="Organ: liver; Vector: pOTB7; Site 1: XhoI; Site 2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library."
 Location/Qualifiers
 BASE COUNT 256 a 279 c 274 g 174 t

Query Match 100.0%; Score 23; DB 14; Length 983;
 Best Local Similarity 100.0%; Pred. No. 5.2;
 Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCAACTGGTGGCCATTGAGCTTC 23
 Db 43 CCAACTGGTGGCCATTGAGCTTC 21

RESULT 13
 BM928196/c 1042 bp mRNA linear EST 12-MAR-2002
 LOCUS AGENCOURT 6715416 NIH_MGC_100 Homo sapiens cDNA clone IMAGE:5797961
 DEFINITION 5', mRNA sequence.
 ACCESSION BM928196
 VERSION BM928196.1 GI:19378575
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens

Query Match 100.0%; Score 23; DB 14; Length 983;
 Best Local Similarity 100.0%; Pred. No. 5.2;
 Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCAACTGGTGGCCATTGAGCTTC 23
 Db 43 CCAACTGGTGGCCATTGAGCTTC 21

RESULT 13
 BM928196/c 1042 bp mRNA linear EST 12-MAR-2002
 LOCUS AGENCOURT 6715416 NIH_MGC_100 Homo sapiens cDNA clone IMAGE:5797961
 DEFINITION 5', mRNA sequence.
 ACCESSION BM928196
 VERSION BM928196.1 GI:19378575
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens

REFERENCE
 AUTHORS NIH-MGC http://mgs.nci.nih.gov/
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: CGAP (Stanford)
 cDNA Library Preparation: Rubin Laboratory
 DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: LLCM2025 row: c column: 18
 High quality sequence stop: 662.

FEATURES
 source
 1..1042
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:5797961"
 /clone_lib="NIH_MGC_100"
 /tissue_type="hepatocellular carcinoma, cell line"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: liver; Vector: pOTB7; Site 1: XhoI; Site 2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library."
 Location/Qualifiers
 BASE COUNT 266 a 299 c 286 g 190 t

Query Match 100.0%; Score 23; DB 14; Length 1042;
 Best Local Similarity 100.0%; Pred. No. 5.3;
 Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCAACTGGTGGCCATTGAGCTTC 23
 Db 43 CCAACTGGTGGCCATTGAGCTTC 21

RESULT 14
 BQ647561/c 1061 bp mRNA linear EST 15-JUL-2002
 LOCUS AGENCOURT 8349614 NIH_MGC_100 Homo sapiens cDNA clone IMAGE:6284767
 DEFINITION 5', mRNA sequence.

ACCESSION BQ647561
 VERSION BQ647561.1 GI:21771733
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 1061)
 NIH-MGC http://mgs.nci.nih.gov/
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: CGAP (Stanford)
 cDNA Library Preparation: Rubin Laboratory
 DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: LLCM2482 row: 0 column: 08
 High quality sequence stop: 443.

Query Match 100.0%; Score 23; DB 14; Length 1061;
 Best Local Similarity 100.0%; Pred. No. 5.3;
 Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCAACTGGTGGCCATTGAGCTTC 23
 Db 43 CCAACTGGTGGCCATTGAGCTTC 21

RESULT 14
 BQ647561/c 1061 bp mRNA linear EST 15-JUL-2002
 LOCUS AGENCOURT 8349614 NIH_MGC_100 Homo sapiens cDNA clone IMAGE:6284767
 DEFINITION 5', mRNA sequence.

ACCESSION BQ647561
 VERSION BQ647561.1 GI:21771733
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 1061)
 NIH-MGC http://mgs.nci.nih.gov/
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: CGAP (Stanford)
 cDNA Library Preparation: Rubin Laboratory
 DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: LLCM2482 row: 0 column: 08
 High quality sequence stop: 443.

FEATURES
source

Location/Qualifiers
 1. .1061
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:6284767"
 /clone_lib="NIH_MGC_100"
 /tissue_type="Hepatocellular carcinoma, cell line"
 /lab_host="DH10B (phage-resistant)"
 /notes="Organ: liver; Vector: pOPB7; Site 1: XhoI; Site 2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library."
 BASE COUNT 277 a 323 c 282 g 179 t
 ORIGIN

Query Match 100.0%; Score 23; DB 14; Length 1061;
 Best Local Similarity 100.0%; Pred. No. 5.4;
 Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 CCAACTGGTGGCCATTGAGCTTC 23
 |||||
 Db 43 CCAACTGGTGGCCATTGAGCTTC 21

RESULT 15
 BE541467/c
 LOCUS 601067919F1 NIH_MGC_10 Homo sapiens cDNA clone IMAGE:3454085 5',
 DEFINITION mRNA sequence.
 ACCESSION BE541467
 VERSION BE541467.1 GI:9770112
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 770)
 NIH-MGC http://mhc.nci.nih.gov/
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: sgapsb@mail.nih.gov
 Tissue Procurement: ATCC

cDNA Library Preparation: Life Technologies, Inc.
 DNA Library Arrayed by: Incyte Genomics, Inc.
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: LHAM8438 row: f column: 06
 High quality sequence stop: 631.

FEATURES
source

Location/Qualifiers
 1. .770
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:3454085"
 /clone_lib="NIH_MGC_10"
 /cell_line="MGC36"
 /lab_host="DH10B"
 /notes="Organ: cervix; Vector: pCMV-SPORT6; Site 1: NotI; Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt. Average insert size 1.5 Kb. Library prepared by Life Technologies."
 BASE COUNT 221 a 210 c 224 g 115 t
 ORIGIN

Query Match 93.0%; Score 21.4; DB 10; Length 770;
 Best Local Similarity 95.7%; Pred. No. 25;

Matches 22; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 CCAACTGGTGGCCATTGAGCTTC 23
 |||||
 Db 305 CCAACTGGTGGCCATTGAGCTTC 283

Search completed: April 18, 2003, 07:31:57
 Job time : 804.985 secs

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 18, 2003, 04:46:26 ; Search time 292 Seconds
(without alignments)
2192.677 Million cell updates/sec

Title: US-09-270-437D-13

Perfect score: 22

Sequence: 1 gctcttgggacaggaagtc 22

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl.*

1: gb.ba.*
2: gb.htg.*
3: gb.in.*
4: gb.om.*
5: gb.ov.*
6: gb.pat.*
7: gb.ph.*
8: gb.pl.*
9: gb.pr.*
10: gb.ro.*
11: gb.sts.*
12: gb.sy.*
13: gb.un.*
14: gb.vi.*
15: em.ba.*
16: em.fun.*
17: em.hum.*
18: em.in.*
19: em.mu.*
20: em.om.*
21: em.or.*
22: em.ov.*
23: em.pat.*
24: em.ph.*
25: em.pl.*
26: em.ro.*
27: em.ste.*
28: em.un.*
29: em.vi.*
30: em.htg.hum.*
31: em.htg.inv.*
32: em.htg.other.*
33: em.htg.mus.*
34: em.htg.pin.*
35: em.htg.rod.*
36: em.htg.mam.*
37: em.htg.vrt.*
38: em.sv.*
39: em.htgo.hum.*
40: em.htgo.mus.*
41: em.htgo.other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	21	95.5	198623	2	AC113611	AC113611 Homo sapi
2	20	90.9	184582	2	AC105949	AC105949 Mus muscu
3	19.4	88.2	2010	9	AF117107	AF117107 Homo sapi
4	19.4	88.2	2935	6	AR059699	AR059699 Sequence
5	19.4	88.2	2935	6	AR083389	AR083389 Sequence
6	19.4	88.2	2935	6	AR164498	AR164498 Sequence
7	19.4	88.2	2935	6	I15470	I15470 Sequence 1
8	19.4	88.2	2935	6	I32484	I32484 Sequence 1
9	19.4	88.2	3283	6	ARI71867	ARI71867 Sequence
10	19.4	88.2	3412	6	ARI71865	ARI71865 Sequence
11	19.4	88.2	3642	9	BC021290	BC021290 Homo sapi
12	19.4	88.2	3667	9	AF057352	AF057352 Homo sapi
13	19.4	88.2	91084	9	AL596177	AL596177 Human DNA
14	19.4	88.2	112259	2	AC026579	AC026579 Homo sapi
15	19.4	88.2	142971	9	AC020629	AC020629 Homo sapi
16	19.4	88.2	146376	9	AC009247	AC009247 Homo sapi
17	19.4	88.2	150595	2	AC095143	AC095143 Rattus no
18	19.4	88.2	181007	2	AC118527	AC118527 Rattus no
19	19.4	88.2	182695	2	AC015706	AC015706 Homo sapi
20	19.4	88.2	193422	2	AC068664	AC068664 Mus muscu
21	19.4	88.2	214592	2	AC084163	AC084163 Mus muscu
22	19.4	88.2	214823	2	AC113869	AC113869 Rattus no
23	19.4	86.4	139351	2	AC129792	AC129792 Rattus no
24	19.4	86.4	167686	2	AC113909	AC113909 Rattus no
25	18.8	85.5	2766	9	AK096182	AK096182 Homo sapi
26	18.8	85.5	7085	8	AF231351	AF231351 Nicotiana
27	18.8	85.5	110000	2	AC021072	Continuation (2 of
28	18.8	85.5	141594	9	AC068302	AC068302 Homo sapi
29	18.8	85.5	190588	2	AC009861	AC009861 Homo sapi
30	18.8	85.5	190610	2	AC130419	AC130419 Homo sapi
31	18.4	83.6	2517	9	AK096200	AK096200 Homo sapi
32	18.4	83.6	57100	2	AC100002	AC100002 Mus muscu
33	18.4	83.6	155301	9	AC041040	AC041040 Homo sapi
34	18.4	83.6	156442	9	AC107385	AC107385 Homo sapi
35	18.4	83.6	174007	8	OSJN00181	OSJN00181 Homo sapi
36	18.4	83.6	186517	9	AC022710	AL662983 Oryza sat
37	18.4	83.6	205998	9	AC113187	AC022710 Homo sapi
38	18.4	83.6	207438	2	AC110819	AC113187 Homo sapi
39	18.4	83.6	237732	2	AC125117	AC110819 Mus muscu
40	18.4	83.6	271144	9	HSXDPB	AC125117 Mus muscu
41	18.4	83.6	330999	2	AC009583	AL590763 Homo sapi
42	18.4	81.8	116304	9	AL139411	AC009583 Homo sapi
43	18.4	81.8	155157	2	AL359549	AL139411 Human DNA
44	18.4	81.8	165524	2	AC129304	AL359549 Homo sapi
45	18.4	81.8	176046	9	AC073366	AC129304 Mus muscu
						AC073366 Homo sapi

ALIGNMENTS

RESULT 1
AC113611
LOCUS
DEFINITION Homo sapiens chromosome 4 clone RP11-421M20, WORKING DRAFT
AC113611
SEQUENCE, 4 unordered pieces.
ACCESSION AC113611.2 GI:20136967
VERSION HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ACTIVEPIN.
KEYWORDS Homo sapiens.
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 198623)
AUTHORS Waterston.R.H.
TITLE The sequence of Homo sapiens clone

Mon Apr 21 10:30:20 2003

Unpublished
2 (bases 1 to 198623)
Waterston,R.H.
Direct Submission
Submitted (04-MAR-2002) Genome Sequencing Center, Washington
University School of Medicine, 444 Forest Park Parkway, St. Louis,
MO 63108, USA
3 (bases 1 to 198623)
Waterston,R.H.
Direct Submission
Submitted (11-APR-2002) Genome Sequencing Center, Washington
University School of Medicine, 444 Forest Park Parkway, St. Louis,
MO 63108, USA
On Apr 11, 2002 this sequence version replaced gi:19073839.

----- Genome Center -----
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu/gsc/index.shtml
Contact: submissions@wustl.edu
----- Project Information -----
Center project name: H.NH0421M20
----- Summary Statistics -----
Sequencing vector: M13; 0%

Chemistry: Dye-primer ET; 0% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 189380 bases at least Q40
Consensus quality: 189735 bases at least Q30
Consensus quality: 199273 bases at least Q20
Insert size: 196000; agarose-fp
Insert size: 206234; sum-of-contigs
Quality coverage: 8.84 in Q20 bases; agarose-fp
Quality coverage: 8.97 in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently
* consists of 4 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1 26542: contig of 26542 bp in length
* 26543: gap of unknown length
* 26543: contig of 30605 bp in length
* 57248: gap of unknown length
* 57347: gap of unknown length
* 108935: contig of 51588 bp in length
* 108936: gap of unknown length
* 109036: contig of 89588 bp in length.

FEATURES

source
1. .198623
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="4"
/clone="RP11-421M20"
misc_feature
1. .26542
/note="assembly_name:Contig12"
misc_feature
26643. .57247
/note="assembly_name:Contig13"
clone end:SP6
vector_side:right
misc_feature
57348. .108935
/note="assembly_name:Contig14"
clone end:T7
vector_side:right
misc_feature
109036. .198623
/note="assembly_name:Contig15"
BASE COUNT 47163 a 54375 g 42946 t 300 others
ORIGIN

Query Match 95.5%; Score 21; DB 2; Length 198623;
Best Local Similarity 100.0%; Pred. No. 4.7;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCTCTTTGGGACAGGAAGT 21
DB 9880 GCTCTTTGGGACAGGAAGT 9900

RESULT 2

AC105949 184582 bp DNA linear HTG 14-APR-2002
Mus musculus clone RP24-217C15, WORKING DRAFT SEQUENCE, 12 ordered
pieces.
AC105949 2 GI:20147936
HTG; HTGS_PHASE2; HTGS_DRAFT; HTGS_FULLTOP.
MUS musculus
MUS musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

REFERENCE

AUTHORS

REFERENCE

AUTHORS

REFERENCE

AUTHORS

REFERENCE

AUTHORS

REFERENCE

AUTHORS

REFERENCE

AUTHORS

REFERENCE

AUTHORS

REFERENCE

AUTHORS

REFERENCE

AUTHORS

REFERENCE

AUTHORS

REFERENCE

AUTHORS

REFERENCE

AUTHORS

REFERENCE

AUTHORS

REFERENCE

AUTHORS

REFERENCE

AUTHORS

REFERENCE

AUTHORS

REFERENCE

AUTHORS

REFERENCE

AUTHORS

REFERENCE

AUTHORS

REFERENCE

TITLE
JOURNAL
COMMENT

Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.
Direct Submission
Submitted (14-APR-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Apr 14, 2002 this sequence version replaced gi:18129448.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: <http://www-seq.wi.mit.edu>
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information -----
Center project name: L18821
Center clone name: 217 C 15

----- Summary Statistics -----
Sequencing vector: Plasmid; n/a; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 181431 bases at least Q40
Consensus quality: 182839 bases at least Q30
Consensus quality: 183305 bases at least Q20
Insert size: 183000; agarose-fp
Insert size: 183482; sum-of-contigs
Quality coverage: 6.9 in Q20 bases; agarose-fp
Quality coverage: 6.9 in Q20 bases; sum-of-contigs

----- NOTE: This is a 'working draft' sequence. It currently
* consists of 12 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* provided by the submitter.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.

1 1013: contig of 1013 bp in length
1014 1113: gap of 100 bp
1114 2223: contig of 1110 bp in length
2224 2323: gap of 100 bp
2324 4327: contig of 2004 bp in length
4328 4427: gap of 100 bp
4428 6625: contig of 2198 bp in length
6626 6725: gap of 100 bp
6726 10232: contig of 3507 bp in length
10233 10332: gap of 100 bp
10333 17743: contig of 7409 bp in length
17742 17841: gap of 100 bp
17842 30472: contig of 12631 bp in length
30473 30572: gap of 100 bp
30573 49894: contig of 19323 bp in length
49895 49994: gap of 100 bp
49995 69993: contig of 19999 bp in length
69994 70093: gap of 100 bp
70094 89184: contig of 19091 bp in length
89185 89284: gap of 100 bp
89285 106186: contig of 16902 bp in length
106187 106286: gap of 100 bp
106287 184582: contig of 78296 bp in length.

FEATURES

source
1. .184582
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone_lib="RP24-217C15"
/clone_lib="RP24-217C15"
1. .1013
/notes="assembly_fragment"
1114. .2223
/notes="assembly_fragment"
2224. .4327
/notes="assembly_fragment"
4428. .6625

misc_feature
1114. .2223
/notes="assembly_fragment"
2224. .4327
/notes="assembly_fragment"
4428. .6625

misc_feature
6726. .10232
/notes="assembly_fragment"
10333. .17741
/notes="assembly_fragment"
17842. .30472
/notes="assembly_fragment"
30573. .49894
/notes="assembly_fragment"
49995. .69993
/notes="assembly_fragment"
70094. .89184
/notes="assembly_fragment"
89285. .106186
/notes="assembly_fragment"
106287. .184582
/notes="assembly_fragment"
BASE COUNT 53205 a 38803 c 37801 g 53673 t 1100 others
ORIGIN

Query Match 90.9%; Score 20; DB 2; Length 184582;
Best Local Similarity 100.0%; Pred. No. 16;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCTCTTTGGGACAGGAAG 20
|||||||
Db 139851 GCTCTTTGGGACAGGAAG 139870

RESULT 3

AF117107
LOCUS
DEFINITION
Homo sapiens IGF-II mRNA-binding protein 2 (IMP-2) mRNA, complete cds.
ACCESSION
AF117107
VERSION
AF117107.1 GI:4191609
SOURCE
Homo sapiens.
ORGANISM
Homo sapiens.

REFERENCE
AUTHORS
Nielsen, J., Christensen, J., Lykke-Andersen, J., Johnsen, A.H.,
Wewer, U.M. and Nielsen, F.C.
1 (bases 1 to 2010)

TITLE
A family of insulin-like growth factor II mRNA-binding proteins
represses translation in late development
Mol. Cell. Biol. 19 (2), 1262-1270 (1999)
JOURNAL
MEDLINE
PUBMED
99108099

REFERENCE
2 (bases 1 to 2010)
Nielsen, J., Christensen, J., Lykke-Andersen, J., Johnsen, A.H.,
Wewer, U.M. and Nielsen, F.C.
Direct Submission

TITLE
JOURNAL
Submitted (30-DEC-1998) Institute of Molecular Biology, University
of Copenhagen, Soelvgade 83H, Copenhagen DK-1307, Denmark
FEATURES
source
1. .2010
/organism="Homo sapiens"
/db_xref="taxon:9606"
1. .2010
/genes="IMP-2"

gene
misc_feature
47
/genes="IMP-2"
/gene="IMP-2"
/note="alternative start-codon gives an additional
N-terminal methionine"

CDS
50. .1846
/genes="IMP-2"
/note="binds multiple sites in IGF-II mRNA 5'UTR"
/codon_start=1
/product="IGF-II mRNA-binding protein 2"
/protein_id="AAD09827.1"
/db_xref="GI:4191610"
/translation="MNKLYIGNLSPAVTADDLRLQFGDKLPLAGQVLLKSGYAFVDY"

us-09-270-437d-13.rge

Mon Apr 21 10:30:20 2003

Best Local Similarity 95.2%; Pred. No. 37; Indels 1; Gaps 0; Mismatches 20; Conservative 0;

QY 2 CTCTTTGGGGACAGGAAGTTC 22

Db 2006 CTCTTTGGGGACAGGAAGTTC 2026

RESULT 6

LOCUS AR164498 2935 bp DNA linear PAT 17-OCT-2001

DEFINITION Sequence 1 from patent US 6274140.

ACCESSION AR164498

VERSION AR164498.1 GI:16237546

KEYWORDS

SOURCE Unknown.

ORGANISM Unclassified.

REFERENCE 1 (bases 1 to 2935)

AUTHORS Jones, S. and Tang, J.

TITLE Calcium independent cytosolic phospholipase A2/B enzymes

JOURNAL Patent: US 6274140-A 1 14-AUG-2001;

FEATURES

source Location/Qualifiers

1..2935

/organism="unknown"

BASE COUNT 697 a 856 c 810 g 572 t

ORIGIN

Query Match 88.2%; Score 19.4; DB 6; Length 2935;

Best Local Similarity 95.2%; Pred. No. 37; Indels 1; Gaps 0; Mismatches 20; Conservative 0;

QY 2 CTCTTTGGGGACAGGAAGTTC 22

Db 2006 CTCTTTGGGGACAGGAAGTTC 2026

RESULT 7

LOCUS I15470 2935 bp DNA linear PAT 02-APR-1996

DEFINITION Sequence 1 from patent US 5465595.

ACCESSION I15470

VERSION I15470.1 GI:1250378

KEYWORDS

SOURCE Unknown.

ORGANISM Unclassified.

REFERENCE 1 (bases 1 to 2935)

AUTHORS Jones, S. and Tang, J.

TITLE Calcium independent cytosolic phospholipase A2/B enzymes

JOURNAL Patent: US 5465595-A 1 14-NOV-1995;

FEATURES

source Location/Qualifiers

1..2935

/organism="unknown"

BASE COUNT 697 a 856 c 810 g 572 t

ORIGIN

Query Match 88.2%; Score 19.4; DB 6; Length 2935;

Best Local Similarity 95.2%; Pred. No. 37; Indels 1; Gaps 0; Mismatches 20; Conservative 0;

QY 2 CTCTTTGGGGACAGGAAGTTC 22

Db 2006 CTCTTTGGGGACAGGAAGTTC 2026

RESULT 8

LOCUS I32484 2935 bp DNA linear PAT 06-FEB-1997

DEFINITION Sequence 1 from patent US 5589170.

ACCESSION I32484

VERSION I32484.1 GI:18232275

KEYWORDS

source Location/Qualifiers

1..2935

/organism="unknown"

BASE COUNT 697 a 856 c 810 g 572 t

ORIGIN

Query Match 88.2%; Score 19.4; DB 6; Length 2935;

Best Local Similarity 95.2%; Pred. No. 37; Indels 1; Gaps 0; Mismatches 20; Conservative 0;

QY 2 CTCTTTGGGGACAGGAAGTTC 22

Db 2006 CTCTTTGGGGACAGGAAGTTC 2026

PDQWAIARAIETLSGKVELHGKIMEVDVSVSKLRSKIOIRNIPPHQWELVDGLIA

QYGTVEQVNTDTETAVVNTVATREAEKAMEKLSGROFENYSKISVIPDEEVS

SPSPQORQGRHSREOCHAPGTSQARQIDFRLRLVPTQFVGALIGREGTINKI

TKQTSRVDIHRKNSGAERKVTTHATPEGTSEACRMILEMOKEADETKLAEIPL

KILANGLVRLIGKEGRNLKKIHEHGTKITISLQDLSIYNPERITTVKGTVEACA

SAEIEIMKGLREAFENDMLAVNQALIPGLNLGALGFSTGLSVLSPAGPRGAPPA

APYHPFTSGISFSSLIYPHQGRGPFPHHSYQEIVNLFIPTQAVGAILIGKGAIK

QLARFAGASIKIAPAEQPPDSERWIIITGPEAQKACQGRIFGKLKEENFPNKEEVK

LEAHIRVPSSTAGRVIGKGTKNLQNLTSAEVILVPRDQPDENEVIVRIIGHFFA

SQPAQRKIREIVQVQKQEQKQPGVAGSQRSK

BASE COUNT 565 a 552 c 532 g 361 t

ORIGIN

Query Match 88.2%; Score 19.4; DB 9; Length 2010;

Best Local Similarity 95.2%; Pred. No. 37; Indels 1; Gaps 0; Mismatches 20; Conservative 0;

QY 1 GCTCTTTGGGGACAGGAAGTTC 21

Db 109 GCTCTTTGGGGACAGGAAGTTC 129

RESULT 4

LOCUS AR059699 2935 bp DNA linear PAT 29-SEP-1999

DEFINITION Sequence 1 from patent US 5840511.

ACCESSION AR059699

VERSION AR059699.1 GI:5986149

KEYWORDS

SOURCE Unknown.

ORGANISM Unclassified.

REFERENCE 1 (bases 1 to 2935)

AUTHORS Jones, S. and Tang, J.

TITLE Methods comprising calcium independent cytosolic phospholipase A/B enzymes

JOURNAL Patent: US 5840511-A 1 24-NOV-1998;

FEATURES

source Location/Qualifiers

1..2935

/organism="unknown"

BASE COUNT 697 a 856 c 810 g 572 t

ORIGIN

Query Match 88.2%; Score 19.4; DB 6; Length 2935;

Best Local Similarity 95.2%; Pred. No. 37; Indels 1; Gaps 0; Mismatches 20; Conservative 0;

QY 2 CTCTTTGGGGACAGGAAGTTC 22

Db 2006 CTCTTTGGGGACAGGAAGTTC 2026

RESULT 5

LOCUS AR083389 2935 bp DNA linear PAT 01-SEP-2000

DEFINITION Sequence 1 from patent US 5976854.

ACCESSION AR083389

VERSION AR083389.1 GI:10010179

KEYWORDS

SOURCE Unknown.

ORGANISM Unclassified.

REFERENCE 1 (bases 1 to 2935)

AUTHORS Jones, S. and Tang, J.

TITLE Calcium independent cytosolic phospholipase A.sub.2 /B enzymes

JOURNAL Patent: US 5976854-A 1 02-NOV-1999;

FEATURES

source Location/Qualifiers

1..2935

/organism="unknown"

BASE COUNT 697 a 856 c 810 g 572 t

ORIGIN

Query Match 88.2%; Score 19.4; DB 6; Length 2935;

Best Local Similarity 95.2%; Pred. No. 37; Indels 1; Gaps 0; Mismatches 20; Conservative 0;

QY 2 CTCTTTGGGGACAGGAAGTTC 22

Db 2006 CTCTTTGGGGACAGGAAGTTC 2026

PDQWAIARAIETLSGKVELHGKIMEVDVSVSKLRSKIOIRNIPPHQWELVDGLIA

QYGTVEQVNTDTETAVVNTVATREAEKAMEKLSGROFENYSKISVIPDEEVS

SPSPQORQGRHSREOCHAPGTSQARQIDFRLRLVPTQFVGALIGREGTINKI

TKQTSRVDIHRKNSGAERKVTTHATPEGTSEACRMILEMOKEADETKLAEIPL

KILANGLVRLIGKEGRNLKKIHEHGTKITISLQDLSIYNPERITTVKGTVEACA

SAEIEIMKGLREAFENDMLAVNQALIPGLNLGALGFSTGLSVLSPAGPRGAPPA

APYHPFTSGISFSSLIYPHQGRGPFPHHSYQEIVNLFIPTQAVGAILIGKGAIK

QLARFAGASIKIAPAEQPPDSERWIIITGPEAQKACQGRIFGKLKEENFPNKEEVK

LEAHIRVPSSTAGRVIGKGTKNLQNLTSAEVILVPRDQPDENEVIVRIIGHFFA

SQPAQRKIREIVQVQKQEQKQPGVAGSQRSK

BASE COUNT 565 a 552 c 532 g 361 t

ORIGIN

Query Match 88.2%; Score 19.4; DB 9; Length 2010;

Best Local Similarity 95.2%; Pred. No. 37; Indels 1; Gaps 0; Mismatches 20; Conservative 0;

QY 1 GCTCTTTGGGGACAGGAAGTTC 21

Db 109 GCTCTTTGGGGACAGGAAGTTC 129

RESULT 4

LOCUS AR059699 2935 bp DNA linear PAT 29-SEP-1999

DEFINITION Sequence 1 from patent US 5840511.

ACCESSION AR059699

VERSION AR059699.1 GI:5986149

KEYWORDS

SOURCE Unknown.

ORGANISM Unclassified.

REFERENCE 1 (bases 1 to 2935)

AUTHORS Jones, S. and Tang, J.

TITLE Methods comprising calcium independent cytosolic phospholipase A/B enzymes

JOURNAL Patent: US 5840511-A 1 24-NOV-1998;

FEATURES

source Location/Qualifiers

1..2935

/organism="unknown"

BASE COUNT 697 a 856 c 810 g 572 t

ORIGIN

Query Match 88.2%; Score 19.4; DB 6; Length 2935;

Best Local Similarity 95.2%; Pred. No. 37; Indels 1; Gaps 0; Mismatches 20; Conservative 0;

QY 2 CTCTTTGGGGACAGGAAGTTC 22

Db 2006 CTCTTTGGGGACAGGAAGTTC 2026

RESULT 5

LOCUS AR083389 2935 bp DNA linear PAT 01-SEP-2000

DEFINITION Sequence 1 from patent US 5976854.

ACCESSION AR083389

VERSION AR083389.1 GI:10010179

KEYWORDS

SOURCE Unknown.

ORGANISM Unclassified.

REFERENCE 1 (bases 1 to 2935)

AUTHORS Jones, S. and Tang, J.

TITLE Calcium independent cytosolic phospholipase A.sub.2 /B enzymes

JOURNAL Patent: US 5976854-A 1 02-NOV-1999;

FEATURES

source Location/Qualifiers

1..2935

/organism="unknown"

BASE COUNT 697 a 856 c 810 g 572 t

ORIGIN

Query Match 88.2%; Score 19.4; DB 6; Length 2935;

Best Local Similarity 95.2%; Pred. No. 37; Indels 1; Gaps 0; Mismatches 20; Conservative 0;

QY 2 CTCTTTGGGGACAGGAAGTTC 22

Db 2006 CTCTTTGGGGACAGGAAGTTC 2026

```

SOURCE      Unknown.
ORGANISM    Unknown.
REFERENCE   1 (bases 1 to 2935)
AUTHORS     Jones, S., and Tang, J.
TITLE       Calcium independent cytosolic phospholipase A.sub.2 /B enzymes
JOURNAL     Patent: US 5589170-A 1 31-DEC-1996;
FEATURES    Location/Qualifiers
            source
            1..2935
            /organism="unknown"
BASE COUNT  697 a 856 c 810 g 572 t
ORIGIN
Query Match      88.2%; Score 19.4; DB 6; Length 2935;
Best Local Similarity 95.2%; Pred. No. 37;
Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy  2  CTCCTTTGGGACAGGAGGTC 22
      |||||
Db  2006 CTCCTGGGACAGGAGGTC 2026

RESULT 9
AR171867
LOCUS      AR171867          3283 bp      DNA      linear      PAT 17-DEC-2001
DEFINITION Sequence 8 from patent US 6297364.
ACCESSION  AR171867
VERSION     AR171867.1 GI:17910817
KEYWORDS   Unknown.
SOURCE     Unknown.
ORGANISM   Unknown.

REFERENCE   1 (bases 1 to 3283)
AUTHORS     Chen, Y.-T., Gure, A., Tsang, S., Stockert, E., Jager, E., Alexander, K.
            and Old, L.J.
TITLE       Isolated nucleic acid molecule encoding cancer associated antigen,
            the antigen itself, and uses thereof
JOURNAL     Patent: US 6297364-A 8 02-OCT-2001;
FEATURES    Location/Qualifiers
            1..3283
            /organism="unknown"
BASE COUNT  945 a 833 c 779 g 725 t 1 others
ORIGIN

Query Match      88.2%; Score 19.4; DB 6; Length 3283;
Best Local Similarity 95.2%; Pred. No. 37;
Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy  1  GCTCTTTGGGACAGGAGGT 21
      |||||
Db  132 GCTCTTTGGGACAGGAGCT 152

RESULT 10
AR171865
LOCUS      AR171865          3412 bp      DNA      linear      PAT 17-DEC-2001
DEFINITION Sequence 6 from patent US 6297364.
ACCESSION  AR171865
VERSION     AR171865.1 GI:17910815
KEYWORDS   Unknown.
SOURCE     Unknown.
ORGANISM   Unknown.

REFERENCE   1 (bases 1 to 3412)
AUTHORS     Chen, Y.-T., Gure, A., Tsang, S., Stockert, E., Jager, E., Alexander, K.
            and Old, L.J.
TITLE       Isolated nucleic acid molecule encoding cancer associated antigen,
            the antigen itself, and uses thereof
JOURNAL     Patent: US 6297364-A 6 02-OCT-2001;
FEATURES    Location/Qualifiers
            1..3412
            /organism="unknown"
BASE COUNT  970 a 887 c 804 g 750 t 1 others
ORIGIN

SOURCE      Unknown.
ORGANISM    Unknown.
REFERENCE   1 (bases 1 to 2935)
AUTHORS     Jones, S., and Tang, J.
TITLE       Calcium independent cytosolic phospholipase A.sub.2 /B enzymes
JOURNAL     Patent: US 5589170-A 1 31-DEC-1996;
FEATURES    Location/Qualifiers
            source
            1..2935
            /organism="unknown"
BASE COUNT  697 a 856 c 810 g 572 t
ORIGIN
Query Match      88.2%; Score 19.4; DB 6; Length 3412;
Best Local Similarity 95.2%; Pred. No. 37;
Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy  1  GCTCTTTGGGACAGGAGGT 21
      |||||
Db  132 GCTCTTTGGGACAGGAGCT 152

RESULT 11
BC021290
LOCUS      BC021290          3642 bp      mRNA      linear      PRI 22-JAN-2002
DEFINITION Homo sapiens, IGF-II mRNA-binding protein 2, clone MGC:29539
            IMAGE:5090334, mRNA, complete cds.
ACCESSION  BC021290
VERSION     BC021290.1 GI:18204200
KEYWORDS   MGC.
SOURCE     Homo sapiens.
ORGANISM   Homo sapiens.

REFERENCE   1 (bases 1 to 3642)
AUTHORS     Strausberg, R.
TITLE       Direct Submission
JOURNAL     Submitted (14-JAN-2002) National Institutes of Health, Mammalian
            Gene Collection (MGC), Cancer Genomics Office, National Cancer
            Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
            USA
            NIH-MGC Project URL: http://mgc.nci.nih.gov
            Contact: MGC help desk
            Email: cgapbs-remail.nih.gov
            Tissue Procurement: ATCC
            CDNA Library Preparation: Rubin Laboratory
            CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: National Institutes of Health Intramural
            Sequencing Center (NISC),
            Gaithersburg, Maryland;
            Web site: http://www.nisc.nih.gov/
            Contact: nisc_mgc@nhgri.nih.gov
            Shevchenko, Y., Wetherby, K.D., Beckstrom-Sternberg, S.M.,
            Benjamin, B., Blakesley, R.W., Bouffard, G.G., Brinkley, C., Brooks, S.,
            Dietrich, N.L., Guan, X., Gupta, J., Ho, S.-L., Karlins, E., Legaspi, R.,
            Lim, M., Maduro, Q.L., Masiello, C., Mastrian, S.D., McCloskey, J.C.,
            McDowell, J., Pearson, R., Snyder, B., Stantripop, S., Thomas, P.J.,
            Tjongson, E.E., Touchman, J.W., Taurgeon, C., Vogt, J.L., Walker, M.A.,
            Zhang, L.-H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAL Plate: 39 Row: 0 Column: 21
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 5729883.

FEATURES    Location/Qualifiers
            1..3642
            /organism="Homo sapiens"
            /db_xref="LocustID:10644"
            /db_xref="taxon:9606"
            /clone="MGC:29539 IMAGE:5090334"
            /tissue_type="Pancreas, epithelioid carcinoma"
            /clone_lib="NIH MGC 42"
            /lab_host="DH10B-R"
            /note="Vector: pOTB7"
            77..1873
            /codon_start=1
            /product="IGF-II mRNA-binding protein 2"
            /protein_id="AAH21290.1"
            /db_xref="GI:18204201"
            /translation="MNKLYIGNLSPAVTADLRDLFGDKLPLAGVLLKSGYAFVDY
            PDQWAIIRAIETLQKVELHGMKINEVDYSYKLSRSKIQIRNPPLHQWELDGLIA
            QYGTVENVEQNTDTETAVNVTYATREAKIWEKLSGHQFENSKISYIDEEVS
            SPSPPQRAQRGDHSRSEQHAPGTSQARQIDFPLRLVPTQFVGAIIIGKEGLTKNI"
            CDS

```

TKQTSRVDIHRKENSRAEKPVTHATPECTSEACRMILEIMOKADETKLAEEIPL
KILANGVLGRILGKGRNLKKIHEHETGKLTISLQDLSTYNERITTVKGTVEACA
SAEIEIMKRLAEFENDMLAVNQANLIPGLNLSALGIFSTGLSVLSPAGPRGAPPA
APYHPTTHSGYFSSLYPHHQPFPFHHSPEQEIYNLPIPTFOAVGANIIGKGAHTK
OLARFAGASIKIAPAEQDVSERWIIITGPPFAKQGRIFGKLEENFENFKEEVK
LEAHVRPSTAGRVIGKGVKNELNLTSAEIVPRDQTPDENEEVIRILIGHFFA
SQTARKIREIVQVQKQEQKYQGVASQRSK"
BASE COUNT 1077 a 919 c 849 g 797 t

Query Match 88.2%; Score 19.4; DB 9; Length 3642;
Best Local Similarity 95.2%; Pred. No. 37;
Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GCTCTTTGGGACAGGAAGT 21
Db 136 GCTCTTTGGGACAGGAAGT 156

RESULT 12
AF057352 3667 bp mRNA linear PRI 23-MAY-1999
LOCUS Homo sapiens hepatocellular carcinoma autoantigen (p62) mRNA,
DEFINITION complete cds.
ACCESSION AF057352
VERSION AF057352.1 GI:4883680
KEYWORDS
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 3667)
Zhang J.Y., Chan, E.K., Peng, X.X. and Tan, E.M.
A novel cytoplasmic protein with RNA-binding motifs is an
autoantigen in human hepatocellular carcinoma
J. Exp. Med. 189 (7), 1101-1110 (1999)
99207072
10190901
2 (bases 1 to 3667)
Zhang, J.Y., Chan, E.K.L., Peng, X.X. and Tan, E.M.
Direct Submission
Submitted (03-APR-1998) Molecular & Experimental Medicine, The
Scripps Research Institute, 10550 N. Torrey Pines Road, La Jolla,
CA 92037, USA

FEATURES
source 1..3667
/organism="Homo sapiens"
/db_xref="taxon:9606"
1..3667
/gene="p62"
436..2106
/gene="p62"
/notes="RNA-binding protein"
/codon_start=1
/product="hepatocellular carcinoma autoantigen"
/protein_id="A0D31596.1"
/db_xref="GI:4883681"
/translations="MMNKLVIIGNLSPAVTDRLRFGDKLPLAGQVLKSGYAFVD
YPDQWAIATLSKGLVEMVDYSVSKLRSRKQIRNPPHLOWELVDGLL
AQGTQVENVQNTDETAVVNTYATREAKINMEKLSGHOFENYSKIPDEEV
SSPSPQAGRHDSREQHAGFGTSQARQIDFPLRLVPTQFVGAIIKGLGTIKN
ITKQTSRVDIHRKENSRAEKPVTHATPECTSEACRMILEIMOKADETKLAEEIP
LKILANGVLGRILGKGRNLKKIHEHETGKLTISLQDLSTYNERITTVKGTVEAC
ASAEIEMKLEAFENDMLAVNQANLIPGLNLSALGIFSTGLSVLSPAGPRGAPPA
TOAVGALIGKGRNLKKIAPAEQDVSERWIIITGPPFAKQGRIFGKLEENFENFKEEVK
GLKLEENFENFKEEVKLEAHVRPSTAGRVIGKGVKNELNLTSAEIVPRDQTP
DENEVIRIIGHFFASQATQAKIREIVQVQKQEQKYQGVASQRSK"
BASE COUNT 1081 a 882 c 846 g 858 t

Query Match 88.2%; Score 19.4; DB 9; Length 3667;
Best Local Similarity 95.2%; Pred. No. 37;
Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GCTCTTTGGGACAGGAAGT 21
Db 498 GCTCTTTGGGACAGGAAGT 518

RESULT 13
AL596177 91084 bp DNA linear PRI 08-SEP-2001
LOCUS Human DNA sequence from clone RP11-325P15 on chromosome 1, complete
DEFINITION sequence.
ACCESSION AL596177
VERSION AL596177.4 GI:15552942
KEYWORDS HTG.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 91084)
Clark, S.
Direct Submission
Submitted (08-SEP-2001) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk
requests: clonerequest@sanger.ac.uk
On Sep 11, 2001 this sequence version replaced gi:15022347.
During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.
This sequence was finished as follows unless otherwise noted: all
regions were either double-stranded or sequenced with an alternate
chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by at least
one plasmid subclone or more than one M13 subclone; and the
assembly was confirmed by restriction digest. The following
abbreviations are used to associate primary accession numbers given
in the feature table with their source databases: Em, EMBL; Sw, SWISSPROT; Tr, TrEMBL; Wp, WORMPEP; Information on the WORMPEP
database can be found at
http://www.sanger.ac.uk/projects/Celegans/wormpep This sequence
was generated from part of bacterial clone contigs of human
chromosome 1, constructed by the Sanger Centre Chromosome 1 Mapping
Group. Further information can be found at
http://www.sanger.ac.uk/HGP/Chr1
RP11-325P15 is from the library RPCI-11.2 constructed by the group
of Pieter de Jong. For further details see
http://www.chori.org/bacpac/home.htm
VECTOR: pBACe3.6
IMPORTANT: This sequence is not the entire insert of clone
RP11-325P15 It may be shorter because we sequence overlapping
sections only once, except for a short overlap.
The true left end of clone RP11-325P15 is at 1 in this sequence.
The true left end of clone RP11-337C18 is at 89085 in this
sequence. The true right end of clone RP4-704D21 is at 41034 in
this sequence.

FEATURES
source 1..91084
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="1"
/clone="RP11-325P15"
/clone_lib="RPCI-11.2"
BASE COUNT 27074 a 19586 c 19377 g 25047 t

Query Match 88.2%; Score 19.4; DB 9; Length 91084;
Best Local Similarity 95.2%; Pred. No. 33;
Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GCTCTTTGGGACAGGAAGT 21

Mon Apr 21 10:30:20 2003

```

* 31202 31301: gap of 100 bp
* 31302 31388: contig of 687 bp in length
* 31389 32088: gap of 100 bp
* 32089 32777: contig of 689 bp in length
* 32778 32877: gap of 100 bp
* 32878 33574: contig of 697 bp in length
* 33575 33674: gap of 100 bp
* 33675 34379: contig of 705 bp in length
* 34380 34479: gap of 100 bp
* 34480 35191: contig of 712 bp in length
* 35192 35291: gap of 100 bp
* 35292 36001: contig of 710 bp in length
* 36002 36101: gap of 100 bp
* 36102 36806: contig of 705 bp in length
* 36807 36906: gap of 100 bp
* 36907 37615: contig of 709 bp in length
* 37616 37715: gap of 100 bp
* 37716 38396: contig of 681 bp in length
* 38397 38496: gap of 100 bp
* 38497 39187: contig of 691 bp in length
* 39188 39287: gap of 100 bp
* 39288 39986: contig of 699 bp in length
* 39987 40086: gap of 100 bp
* 40087 40788: contig of 702 bp in length
* 40789 40888: gap of 100 bp
* 40889 41603: contig of 715 bp in length
* 41604 41703: gap of 100 bp
* 41704 42424: contig of 721 bp in length
* 42425 42524: gap of 100 bp
* 42525 43231: contig of 707 bp in length
* 43232 43331: gap of 100 bp
* 43332 44029: contig of 698 bp in length
* 44030 44129: gap of 100 bp
* 44130 44815: contig of 686 bp in length
* 44816 44915: gap of 100 bp
* 44916 45611: contig of 696 bp in length
* 45612 45711: gap of 100 bp
* 45712 46403: contig of 692 bp in length
* 46404 46503: gap of 100 bp
* 46504 47206: contig of 703 bp in length
* 47207 47306: gap of 100 bp
* 47307 47995: contig of 689 bp in length
* 47996 48095: gap of 100 bp
* 48096 48793: contig of 698 bp in length
* 48794 48893: gap of 100 bp
* 48894 49463: contig of 570 bp in length
* 49464 49563: gap of 100 bp
* 49564 50250: contig of 687 bp in length
* 50251 50350: gap of 100 bp
* 50351 51091: contig of 741 bp in length
* 51092 51191: gap of 100 bp
* 51192 51909: contig of 718 bp in length
* 51910 52009: gap of 100 bp
* 52010 52713: contig of 704 bp in length
* 52714 52813: gap of 100 bp
* 52814 53526: contig of 713 bp in length
* 53527 53626: gap of 100 bp
* 53627 54337: contig of 711 bp in length
* 54338 54437: gap of 100 bp
* 54438 55155: contig of 718 bp in length

```

Query Match 88.2%; Score 19.4; DB 2; Length 112259;
 Best Local Similarity 95.2%; Pred. No. 33;
 Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GCTCTTTGGGACAGGAGGT 21
 DB 77783 GCTCTTTGGGACAGGAGGT 77763

RESULT 15
 AC020629
 LOCUS AC020629 142971 bp DNA linear PRI 07-MAR-2002

DEFINITION

Human BAC Library) complete sequence.
 AC020629 GI:7656675
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM

Homo sapiens

Homo sapiens

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 142971)

Muzny,D.M., Adams,C., Bailey,M., Barbara,J., Blankenburg,K.,

Bodota,B., Bouck,J., Bowie,S., Brooks,A., Bunay,C., Bunac,C.,

Burkett,C., Burrows,J., Carter,M., Chacko,J., Chen,Z., Cox,C.,

David,R., Delgado,O., Deshazo,D., Ding,Y., Domah-Rashid,N.,

Dugan-Rocha,S., Durbin,K.J., Fernandez,C., Ferraguto,D.,

Guevara,W., Harris,K., Hernandez,J., Hodgson,A., Hogues,M.,

Holloway,C., Hosak,H., Issat,A., Jackson,L.E., Jackson,L., Jia,Y.,

Jones,M., Kelly,S., Kneitz,S., Kondajewski,N., Kong,Y., Kovar,C.,

Lau,S., Leal,B., Lee,E., Li,Z., Lichtarge,O., Liu,J., Liu,W.,

Logan,O., Lozano,R.J., Lu,J., Lucier,R., Marondel,I., Martin,R.,

Martinez,C., McLeod,M.P., Mei,G., Merscher,S., Miller,A.,

Montgomery,K.T., Morgan,M., Morris,S., Nash,S., Nelson,A.,

Nguyen,R., Nguyen,N., Nguyen,S., Osawai,G., Parish,B., Paxton,S.,

Payton,B., Perez,L., Pu,L.L., Quiles,M., Reiter,D., Rives,M.,

Samuel,S., Say,J., Scherer,S., Shah,E., Shen,H., Shim,C., Simon,M.,

Sparks,A., Stamps,A., Sugang,R., Tabor,F., Taylor,T., Vasquez,L.,

Vinson,R., Vo,Q., Wabwah,M., Watlington,S., Weinstein,G.,

Weinstock,I.R., Williamson,A., Worley,K., Wren,J., Wrensford,G.,

Xhang,A.M., Yang,R., Yu,W., Zhou,X., Kucherlapati,R., Nelson,D. and

Gibbs,R.

Direct Submission

Unpublished

2 (bases 1 to 142971)

Worley,K.C.

Direct Submission

Submitted (07-JAN-2000) Human Genome Sequencing Center, Department

of Molecular and Human Genetics, Baylor College of Medicine, One

Baylor Plaza, Houston, TX 77030, USA

3 (bases 1 to 142971)

Worley,K.C.

Direct Submission

Submitted (27-APR-2000) Human Genome Sequencing Center, Department

of Molecular and Human Genetics, Baylor College of Medicine, One

Baylor Plaza, Houston, TX 77030, USA

4 (bases 1 to 142971)

Worley,K.C.

Direct Submission

Submitted (28-APR-2000) Human Genome Sequencing Center, Department

of Molecular and Human Genetics, Baylor College of Medicine, One

Baylor Plaza, Houston, TX 77030, USA

5 (bases 1 to 142971)

Worley,K.C.

Direct Submission

Submitted (07-MAR-2002) Human Genome Sequencing Center, Department

of Molecular and Human Genetics, Baylor College of Medicine, One

Baylor Plaza, Houston, TX 77030, USA

On Apr 27, 2000 this sequence version replaced gi:7025656.

INFORMATION: <http://www.hgsc.bcm.tmc.edu/> or email

gc-help@bcm.tmc.edu

Homo sapiens 12q BAC RP11-76E16 (Roswell Park Cancer Institute

Human BAC Library) complete sequence.

AC020629 GI:7656675

VERSION

KEYWORDS

SOURCE

ORGANISM

Homo sapiens

Homo sapiens

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 142971)

Muzny,D.M., Adams,C., Bailey,M., Barbara,J., Blankenburg,K.,

Bodota,B., Bouck,J., Bowie,S., Brooks,A., Bunay,C., Bunac,C.,

Burkett,C., Burrows,J., Carter,M., Chacko,J., Chen,Z., Cox,C.,

David,R., Delgado,O., Deshazo,D., Ding,Y., Domah-Rashid,N.,

Dugan-Rocha,S., Durbin,K.J., Fernandez,C., Ferraguto,D.,

Guevara,W., Harris,K., Hernandez,J., Hodgson,A., Hogues,M.,

Holloway,C., Hosak,H., Issat,A., Jackson,L.E., Jackson,L., Jia,Y.,

Jones,M., Kelly,S., Kneitz,S., Kondajewski,N., Kong,Y., Kovar,C.,

Lau,S., Leal,B., Lee,E., Li,Z., Lichtarge,O., Liu,J., Liu,W.,

Logan,O., Lozano,R.J., Lu,J., Lucier,R., Marondel,I., Martin,R.,

Martinez,C., McLeod,M.P., Mei,G., Merscher,S., Miller,A.,

Montgomery,K.T., Morgan,M., Morris,S., Nash,S., Nelson,A.,

Nguyen,R., Nguyen,N., Nguyen,S., Osawai,G., Parish,B., Paxton,S.,

Payton,B., Perez,L., Pu,L.L., Quiles,M., Reiter,D., Rives,M.,

Samuel,S., Say,J., Scherer,S., Shah,E., Shen,H., Shim,C., Simon,M.,

Sparks,A., Stamps,A., Sugang,R., Tabor,F., Taylor,T., Vasquez,L.,

Vinson,R., Vo,Q., Wabwah,M., Watlington,S., Weinstein,G.,

Weinstock,I.R., Williamson,A., Worley,K., Wren,J., Wrensford,G.,

Xhang,A.M., Yang,R., Yu,W., Zhou,X., Kucherlapati,R., Nelson,D. and

Gibbs,R.

Direct Submission

Unpublished

2 (bases 1 to 142971)

Worley,K.C.

Direct Submission

Submitted (07-JAN-2000) Human Genome Sequencing Center, Department

of Molecular and Human Genetics, Baylor College of Medicine, One

Baylor Plaza, Houston, TX 77030, USA

3 (bases 1 to 142971)

Worley,K.C.

Direct Submission

Submitted (27-APR-2000) Human Genome Sequencing Center, Department

of Molecular and Human Genetics, Baylor College of Medicine, One

Baylor Plaza, Houston, TX 77030, USA

4 (bases 1 to 142971)

Worley,K.C.

Direct Submission

Submitted (28-APR-2000) Human Genome Sequencing Center, Department

of Molecular and Human Genetics, Baylor College of Medicine, One

Baylor Plaza, Houston, TX 77030, USA

5 (bases 1 to 142971)

Worley,K.C.

Direct Submission

Submitted (07-MAR-2002) Human Genome Sequencing Center, Department

of Molecular and Human Genetics, Baylor College of Medicine, One

Baylor Plaza, Houston, TX 77030, USA

On Apr 27, 2000 this sequence version replaced gi:7025656.

INFORMATION: <http://www.hgsc.bcm.tmc.edu/> or email

CLONE LENGTH: This sequence does not necessarily represent the entire insert of this clone. Overlapping regions of clones are only sequenced and submitted once, so the sequence for the remainder of the insert may be found in the record for the adjacent clones. Overlapping clones are noted at the beginning and end of the Features listing.

ANNOTATION OF FEATURES:
 STSs are identified using ePCR (Genome Res. 7:541-550) searches of a local database that includes entries from dbSTS, GDB, and local mapping efforts.
 Repeats are identified using RepeatMasker (A. Smit and P. Green,

unpublished.) for Human and Mouse sequences.

Genes and Region of sequence similarity are identified by BLAST (Nuc. Acids Res 25:3389-3402) similarity (expect < 1e-34) to the EST and cDNA sequences. Genes demonstrate at least two exons flanked by consensus splice sites that maintained sequence continuity across the splice junctions. Sequences that are not identical matches are annotated as similar.

SEQUENCING READ COVERAGE: Sequencing is completed to a minimum standard of double strand coverage with a minimum of 2 clones and 2 reads with no ambiguities or 2 chemistries with a minimum of 2 clones and 3 reads with no ambiguities. If the sequence quality for a region does not meet this standard, it will be indicated in the annotation as Low Coverage.

QUALITY OF INDIVIDUAL BASES: This sequence meets stringent quality standards - estimated error rate less than 1 per 10,000 bases. Reports of lowest quality individual bases and measures of base quality are listed below. Description of the metrics can be found at URL: <http://gc.bcm.tmc.edu:8088/quality.info/genbank.annotation.html>.

QUALSTAT-REPORT-----

```

----- Summary Statistics -----
Contig length: 142971
Phrap values in estimate: 140680
Average error rate (BCM-Phrap estimate): 0.00018703
Fraction of Phrap values less than 40: 0.0433608
Number of consensus changing edits: 31
Number of N's in consensus: 0

```

```

----- Consensus changing edits -----
Position Original+Context Edited+Context
11171 gtctctgga(n)caaaaaaaa gtctctgga(a)caaaaaaaa
12018 tatctatcg(n)caatatatat tatctatcg(t)ctatcttat
16604 tccactctaa(n)agttccattt tccactctaa(g)agttccattt
18395 tagtaaccac(n)tacattttta tagtaaccac(t)tacattttta
21879 aaaaaaaa(n)nttaaaaca aaaaaaaa(a)attaaacag
21880 aaaaaaaa(n)nttaaaacag aaaaaaaa(a)attaaacag
21881 aaaaaaaa(n)nttaaaacag aaaaaaaa(a)attaaacag
21906 cctaggtccc(n)cttaaatgag cctaggtccc(a)cttaaatgag
22169 accagctcg(n)cttaaatgag accagctcg(c)caacatggcg
38915 ttacatctag(n)ttcttgcaaa ttacatctag(t)ttcttgcaaa
55675 acttggcttc(n)aaaaaaaaa acttggcttc(a)aaaaaaaaa
57558 gaccagctg(n)ccacatgggt gaccagctg(a)ccacatgggt
74371 gtctcgccc(n)ccgcccgcgc gtctcgccc(c)ccgcccgcgc
88248 taatgaggg(n)agaccagtgg taatgaggg(c)agaccagtgg
94121 gcagtggcc(n)agaccagcc gcagtggcc(c)agaccagcc
95228 ttgtctcga(n)gcctctctgc ttgtctcga(t)gcctctctgc
95230 gtcacatgt(n)nnnttgtaa gtcacatgt(c)nnnttgtaa
95231 catgtgtn(n)nnnttgtaa catgtgtn(c)nnnttgtaa
95232 agtgnn(n)ttgttcaact agtgnn(t)ttgttcaact
104931 ttttttttag(n)acagacaggg ttttttttag(t)acagacaggg
118168 gccacgtct(n)ccaaagtgt gccacgtct(c)ccaaagtgt
120974 catctcaaa(n)anangaac catctcaaa(a)anangaac
120976 tctcaaaa(n)anangaact tctcaaaa(a)anangaact
120978 tcaaaaana(n)angaactct tcaaaaana(a)angaactct
120980 aaaaana(n)gaactctctc aaaaana(a)gaactctctc
139505 ccagccatgg(n)ggngggcacc ccagccatgg(t)ggngggcacc
139508 gccatgng(n)ggngggcacc gccatgng(t)ggngggcacc
139698 attaccatta(n)aatatggtat attaccatta(t)aatatggtat
140104 accgtgccca(n)gagatagaa accgtgccca(g)gagatagaa
140149 tggacaact(n)gataccaca tggacaact(g)gataccaca

```

----- Distribution of Quality < 40 Bases -----

```

1000|
900|

```

```

#
bases
800
700
600
500
400
300
200
100
0

```

```

Version: 1.01 qxfo.
FEATURES
Source
1. .142971 Location/Qualifiers
/db_xref="taxon:9606"
/chromosome="12"
/clone="RP11-76E16"
147. .236
/rpt_family="A-rich"
repeat_region
147. .236
/rpt_family="A-rich"
repeat_region
258. .378
/rpt_family="AluJb"
complement(537. .840)
/rpt_family="AluSx"
repeat_region
complement(1722. .2137)
/rpt_family="nigger2a"
repeat_region
complement(3161. .3584)
/rpt_family="MSTD"
3228. .3502
/standard_name="G05956"
/db_xref="dbSTS:11453"
3347. .3490
/standard_name="Cda0y/b04"
/db_xref="GDB:446196"
3585. .3735
/note="Region similar to Hs#S936143 aj27d02.s1 Homo sapiens cDNA: AA789222"
3885. .3910
/rpt_family="(TA)n"
repeat_region
complement(3917. .4030)
/rpt_family="AluJb"
repeat_region
complement(4544. .4681)
/rpt_family="MIR"
5145. .5510
/note="Region similar to Hs#S953417 oa06g05.s1 Homo sapiens cDNA: AA765439"
5553. .5859
/note="Region similar to Hs#S1195593 qt69f10.x1 Homo sapiens cDNA: A1285403"

```

```

Query Match 88.2%; Score 19.4; DB 9; Length 142971;
Best Local Similarity 95.2%; Pred. No. 32;
Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

Oy 1 GCTCTTTGGGACAGGAAGCT 21

Db 133084 GCTCTTTGGGACAGGAAGCT 133104

Search completed: April 18, 2003, 06:27:37
Job time : 414 secs

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 18, 2003, 04:12:35 / Search time 98 Seconds
(without alignments)
505.551 Million cell updates/sec

Title: US-09-270-437D-13

Perfect score: 22

Sequence: 1 gctcttggggacaggagtc 22

Scoring table: IDENTITY NUC

Gapop 10.0, Gapext 1.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :			
N_Geneseq 101002.*			
1:	/SID22/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.*		
2:	/SID22/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.*		
3:	/SID22/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.*		
4:	/SID22/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.*		
5:	/SID22/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.*		
6:	/SID22/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.*		
7:	/SID22/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.*		
8:	/SID22/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.*		
9:	/SID22/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.*		
10:	/SID22/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.*		
11:	/SID22/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.*		
12:	/SID22/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.*		
13:	/SID22/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.*		
14:	/SID22/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.*		
15:	/SID22/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.*		
16:	/SID22/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.*		
17:	/SID22/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.*		
18:	/SID22/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.*		
19:	/SID22/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.*		
20:	/SID22/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.*		
21:	/SID22/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.*		
22:	/SID22/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.*		
23:	/SID22/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.*		
24:	/SID22/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.*		

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	22	100.0	22	21	AAZ36159
2	19.4	88.2	364	21	AAZ36159
3	19.4	88.2	2010	23	AAZ36159
4	19.4	88.2	2335	17	AAZ36159
5	19.4	88.2	2335	17	AAZ36159
6	19.4	88.2	2335	18	AAZ36159
7	19.4	88.2	2335	18	AAZ36159
8	19.4	88.2	2335	20	AAZ36159
9	19.4	88.2	3283	21	AAZ36159

10	19.4	88.2	3412	21	AAZ36152
11	19.4	88.2	3667	23	AAZ36152
12	19.4	88.2	3694	22	AAZ36150
13	17.8	80.9	1212	19	AAZ36150
14	17.8	80.9	11849	20	AAZ36150
15	17.4	79.1	3036	22	AAZ36150
16	17.2	78.2	429	24	AAZ36150
17	17.2	78.2	778	22	AAZ36150
18	17.2	78.2	801	22	AAZ36150
19	17.2	78.2	4985	20	AAZ36150
20	17.2	78.2	5526	23	AAZ36150
21	17.2	78.2	5778	22	AAZ36150
22	17.2	78.2	5804	22	AAZ36150
23	17.2	78.2	6042	23	AAZ36150
24	16.8	76.4	401	22	AAZ36150
25	16.8	76.4	401	22	AAZ36150
26	16.8	76.4	401	22	AAZ36150
27	16.8	76.4	1442	21	AAZ36150
28	16.8	76.4	4359	23	AAZ36150
29	16.8	76.4	4394	13	AAZ36150
30	16.8	76.4	5364	22	AAZ36150
31	16.8	76.4	5364	22	AAZ36150
32	16.4	74.5	544	22	AAZ36150
33	16.4	74.5	927	19	AAZ36150
34	16.4	74.5	952	22	AAZ36150
35	16.4	74.5	952	22	AAZ36150
36	16.4	74.5	1452	21	AAZ36150
37	16.4	74.5	1452	21	AAZ36150
38	16.4	74.5	1452	24	AAZ36150
39	16.4	74.5	1487	21	AAZ36150
40	16.4	74.5	1599	24	AAZ36150
41	16.4	74.5	1680	18	AAZ36150
42	16.4	74.5	1680	19	AAZ36150
43	16.4	74.5	1680	19	AAZ36150
44	16.4	74.5	1680	20	AAZ36150
45	16.4	74.5	1680	20	AAZ36150

ALIGNMENTS

RESULT 1
AAZ36159
ID AAZ36159 standard; DNA; 22 BP.
XX AAZ36159;
AC AAZ36159;
XX AAZ36159;
DT 11-FEB-2000 (first entry)
DE PCR primer for DNA encoding cancer associated antigen KOC-3.
XX PCR primer for DNA encoding cancer associated antigen KOC-3.
XX Cancer associated antigen; KOC-1; cancer; vaccine; CT7; PCR primer; ss.
XX Synthetic.
XX Homo sapiens.
XX WO9954738-A1.
XX 28-OCT-1999.
XX 16-MAR-1999; 99WO-US05766.
XX 17-APR-1998; 98US-0061709.
XX (LUDW-) LUDWIG INST CANCER RES.
XX Chen Y, Gure A, Tsang S, Stockert E, Jager E, Knuth A, Old LJ;
XX WPI; 2000-013284/01.
XX Nucleotides representing cancer-associated genes, used to develop
XX products for the diagnosis, monitoring and treatment of cancers -

The present sequence is one of a large number of 5' ESTs derived from mRNAs encoding secreted proteins. An ORF has been identified within the sequence. The 5' ESTs were prepared from total human RNAs or poly(A)⁺ RNAs derived from 30 different tissues. EST sequences usually correspond mainly to the 3' untranslated region (UTR) of the mRNA because they are often obtained from oligo-dT primed cDNA libraries. Such ESTs are not well suited for isolating cDNA sequences derived from the 5' ends of mRNAs and even in those cases where longer cDNA sequences have been obtained, the full 5' UTR is rarely included. 5' ESTs are derived from

CC Note: The sequence data for this patent did not appear in the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 2010 BP; 565 A; 552 C; 532 G; 361 T; 0 other;

Query Match 88.2%; Score 19.4; DB 23; Length 2010;
 Best Local Similarity 95.2%; Pred. No. 13;
 Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GCTCTTTGGGACAGGAAGGT 21
 Db 109 GCTCTTTGGGACAGGAAGCT 129

RESULT 4
 AAT44578
 ID AAT44578 standard; cDNA; 2935 BP.
 AC AAT44578;
 XX
 DT 11-FEB-1997 (first entry)
 XX
 DE Calcium-independent cytosolic phospholipase A2/B coding sequence.
 XX
 KW cPLA2/B; calcium-independent cytosolic phospholipase A2/B;
 KW arachidonic acid; 1-palmitoyl-2-[14C]-arachidonyl-phosphatidylcholine;
 KW anti-inflammatory; screen; rheumatoid arthritis; ds.
 XX
 OS Synthetic.

XX
 XX Key Location/Qualifiers
 FT CDS 96..2354
 FT /*tag= a
 XX
 PN US5554511-A.
 PD 10-SEP-1996.
 XX
 XX 27-JUL-1994; 94US-0281193.
 XX
 PR 27-JUL-1994; 94US-0281193.
 PR 14-APR-1995; 95US-0422420.
 XX
 PA (GEMY) GENETICS INST INC.
 XX
 PI Jones S, Tang J;
 XX
 DR WPI; 1996-424653/42.
 DR P-PSDB; AAW01479.

XX
 XX Prodn. of a Ca-independent cytosolic phospholipase A(2)/B - by
 XX culturing host cells contg. the phospholipase gene, useful for
 XX screening anti-inflammatory agents for treating e.g. rheumatoid
 XX arthritis
 XX
 PS Claim 1; Column 15-22; 24pp; English.
 XX
 CC The present sequence encodes a calcium-independent cytosolic
 CC phospholipase A2/B (cPLA2/B) enzyme. cPLA2/B is important in the release
 CC of arachidonic acid in specific tissues characterised by unique membrane
 CC phospholipids. The invention provides a process for producing such an
 CC enzyme. The enzyme has a mol. wt. of 86 kD on SDS-PAGE and the presence
 CC of one or more amino acid sequences selected from AAW01480-92. cPLA2/B
 CC has activity in a mixed micelle assay with
 CC 1-palmitoyl-2-[14C]-arachidonyl- phosphatidylcholine. The enzyme is
 CC useful for screening anti-inflammatory agents mediated by the arachidonic
 CC acid cascade, for treating, e.g. rheumatoid arthritis
 XX
 SQ Sequence 2935 BP; 697 A; 856 C; 810 G; 572 T; 0 other;

Query Match 88.2%; Score 19.4; DB 17; Length 2935;
 Best Local Similarity 95.2%; Pred. No. 14;

Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 Qy 2 CTCTTTGGGACAGGAAGTTC 22
 Db 2006 CTCTTTGGGACAGGAAGTTC 2026

RESULT 5
 AAT05842
 ID AAT05842 standard; cDNA; 2935 BP.
 XX
 AC AAT05842;
 XX

DT 15-JUN-1996 (first entry)
 XX
 DE Calcium-independent cytosolic phospholipase-A2/B enzyme cDNA.
 XX
 KW CHO; calcium-independent cytosolic phospholipase-A2/B; enzyme;
 KW phospholipase-A2; phospholipase-B; drug screening; antiinflammatory;
 KW antibody; ds.
 XX
 OS Cricetulus griseus.

XX
 XX Key Location/Qualifiers
 FT CDS 96..2352
 FT /*tag= a
 XX
 PN US5466595-A.
 XX
 PD 14-NOV-1995.
 XX

XX
 PF 27-JUL-1994; 94US-0281193.
 XX
 PR 27-JUL-1994; 94US-0281193.
 XX

PA (GEMY) GENETICS INST INC.
 XX
 PI Jones S, Tang J;
 XX
 DR WPI; 1996-009526/01.
 DR P-PSDB; AAR93018.

XX
 XX Isolated polynucleotide encoding cytosolic phospholipase A2/8 - for
 XX producing enzyme for use in screening anti-inflammatory agents and
 XX prodn. of antibodies
 XX
 PS Claim 1; Column 15-22; 24pp; English.

XX
 CC This cDNA was prepared from RNA isolated from CHO-DUX cells. The
 CC cDNA may be ligated into predigested lambda phage ZapII/EcoRI vector
 CC and packaged into phage particles. The cDNA library was screened
 CC using oligonucleotide probes based on peptide sequences of calcium-
 CC independent phospholipase-A2/B. This polynucleotide may be
 CC operably linked to an expression control sequence such as pMT2
 CC or pED expression vectors in order to produce the phospholipase
 CC enzyme recombinantly. Host cells include animal cells, insect
 CC cells, eukaryotes, prokaryotes, etc. The sequence may also be
 CC expressed in transgenic animals (e.g. milk of transgenic cow).
 CC The encoded protein is used to screen for agents which inhibit
 CC phospholipase activity for use as antiinflammatory agents.

XX
 SQ Sequence 2935 BP; 697 A; 856 C; 810 G; 572 T; 0 other;

Query Match 88.2%; Score 19.4; DB 17; Length 2935;
 Best Local Similarity 95.2%; Pred. No. 14;
 Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 CTCTTTGGGACAGGAAGTTC 22
 Db 2006 CTCTTTGGGACAGGAAGTTC 2026

RESULT 6

us-09-270-437d-13.rng

Mon Apr 21 10:30:22 2003

AAT59199
ID AAT59199 standard; cDNA; 2935 BP.
AC AAT59199;
XX
XX
XX 20-MAY-1997 (first entry)
DT
XX
XX Ca-independent phospholipase A2/B gene.
DE
XX Ca-independent phospholipase A2/B; Chinese hamster ovary cell; probe;
KW chromatography; DEAE anion exchange; hydrophobic interaction; lambda;
KW heparin Toyopearl; chromatofocusing; eukaryotic expression vector; COS;
KW CHO; inhibitor; anti-inflammatory agent; arachidonic acid cascade; ds.
XX
XX Chinese hamster ovary cells.
OS
XX
XX
XX Key Location/Qualifiers
FH 96..2354
FT CDS /*tag= a
FT /product= ca-independent phospholipase A2/B
FT
XX
XX US5589170-A.
XX
XX 31-DEC-1996.
XX
XX 27-JUL-1994; 94US-0281193.
XX
XX 27-JUL-1994; 94US-0281193.
XX
XX 14-APR-1995; 95US-0422106.
XX
XX (GEMY) GENETICS INST INC.
XX
XX Jones S, Tang J;
XX WPI; 1997-076789/07.
XX P-PSDB; AAW13163.
XX
XX Compn. comprising calcium-independent phospholipase enzyme - for
PT screening for anti-inflammatory agents
PT
XX Claim 5; Column 15-22; 24pp; English.
XX
XX This is the nucleotide sequence encoding the Ca-independent
XX phospholipase A2/B from Chinese hamster ovary cells. The protein was
XX isolated from these cells by conventional chromatographic methods e.g.
XX DEAE anion exchange, hydrophobic interaction, heparin Toyopearl and Mono
XX P 5/20 chromatofocusing chromatography. The purified protein has mol.
XX wt. of 86 kD and an optimum pH 6. The protein was used for amino acid
XX sequencing from which pools of degenerate probes were synthesised. The
XX probes were used to screen a CHO cell cDNA library in lambda ZAPII
XX vector. Of 40000 recombinant phages screened, 12 positive plaques were
XX isolated. One of these, designated clone 9, contained this sequence.
XX The phospholipase gene can be inserted into eukaryotic vectors for
XX expression in COS or CHO cells. The protein, or peptides derived from it
XX e.g. AAW13164-76, can be used to identify phospholipase inhibitors that
XX can be used as anti-inflammatory agents, esp. against components of the
XX arachidonic acid cascade.
XX
XX Sequence 2935 BP; 697 A; 856 C; 810 G; 572 T; 0 other;
SQ
Query Match 88.2%; Score 19.4; DB 18; Length 2935;
Best Local Similarity 95.2%; Pred. No. 14;
Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 2 CTCTTTGGGACAGGAGGTC 22
Db 2006 CTCTCTGGGACAGGAGGTC 2026
RESULT 7
AAT68827
ID AAT68827 standard; cDNA; 2935 BP.
XX

AAT68827;
AC
XX
XX 07-AUG-1997 (first entry)
XX
XX Hamster cytosolic phospholipase A2/B cDNA.
DE
XX Cytosolic phospholipase A2/B; sPLA2/B; arachidonic acid cascade;
KW inflammation; inhibitor; antiinflammatory; CHO; ds.
XX
XX Cricetus sp.
OS
XX Location/Qualifiers
FH Key 96..2354
FT CDS /*tag= a
FT
XX WO9717448-A2.
XX
XX 15-MAY-1997.
XX
XX 07-NOV-1996; 96WO-US17794.
XX
XX 08-NOV-1995; 95US-0555568.
XX
XX (GEMY) GENETICS INST INC.
XX
XX Jones S, Tang J;
XX WPI; 1997-281037/25.
XX P-PSDB; AAW17849.
XX
XX Calcium independent phospholipase A2/B - used to reduce inflammation
PT in a mammalian subject
PT
XX Example 4; Page 33-36; 74pp; English.
XX
XX A cDNA clone (AAT68827), deposited as ATCC 69669, codes for a
XX cytosolic phospholipase A2/B (sPLA2/B) (AAW17849) that is thought to
XX be involved in the arachidonic acid cascade. It was obtd. by
XX screening a CHO-DUX cDNA library with probes based on isolated
XX peptides of the enzyme. Recombinant sPLA2/B was expressed in
XX COS and CHO cells. Probes based on the hamster sPLA2/B clone
XX were used to identify human sPLA2/B partial cDNA clones (see also
XX AAT68823-26).
XX
XX Sequence 2935 BP; 697 A; 856 C; 810 G; 572 T; 0 other;
SQ
Query Match 88.2%; Score 19.4; DB 18; Length 2935;
Best Local Similarity 95.2%; Pred. No. 14;
Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 2 CTCTTTGGGACAGGAGGTC 22
Db 2006 CTCTCTGGGACAGGAGGTC 2026
RESULT 8
AAT64840
ID AAT64840 standard; cDNA; 2935 BP.
XX
XX AAT64840;
XX
XX 02-FEB-1999 (first entry)
XX
XX Chinese hamster calcium independent cPLA2/B cDNA.
DE
XX Calcium independent; cytosolic phospholipase A2/B; screening;
KW anti-inflammatory; arachidonic acid cascade; Chinese hamster; ds.
XX
XX Cricetus griseus.
XX
XX Key Location/Qualifiers
FH Key 96..2354
FT CDS /*tag= a
FT

FT XX US5840511-A. /product= "calcium independent cPLA2/B"
 PN XX 24-NOV-1998.
 PD XX 23-OCT-1996; 96US-0735716.
 PF XX 27-JUL-1994; 94US-0281193.
 PR XX 14-APR-1995; 95US-0422106.
 PS XX 23-OCT-1996; 96US-0735716.
 PT (GEMV) GENETICS INST INC.
 PI Jones S, Tang J;
 XX WPI; 1999-034032/03.
 DR P-PSDB; AAW81825.

XX Screening assay for phospholipase inhibitors - using specified
 PT phospholipase polypeptide
 XX Claim 1b(1); Column 15-22; 24pp; English.

XX This sequence encodes a novel calcium independent cytosolic phospholipase
 CC A2/B enzyme isolated from chinese hamster ovary cells. This protein can
 CC be used for screening unknown compounds for anti-inflammatory activity
 CC mediated by the arachidonic acid cascade.

XX Sequence 2935 BP; 697 A; 856 C; 810 G; 572 T; 0 other;

Query Match 88.2%; Score 19.4; DB 20; Length 2935;
 Best Local Similarity 95.2%; Pred. No. 14;
 Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 CTCTTTGGGGACAGGAGTC 22
 |||||
 Db 2006 CTCTTTGGGGACAGGAGTC 2026

RESULT 9
 AA236154
 ID AA236154 standard; DNA; 3283 BP.
 AC AA236154;
 XX 11-FEB-2000 (first entry)
 DT An alternative form of DNA encoding cancer associated antigen KOC-3.
 DE Cancer associated antigen; KOC-3; cancer; vaccine; CT7; ss.

XX Homo sapiens.
 XX WO9954738-A1.
 XX 28-OCT-1999.
 PD 16-MAR-1999; 99WO-US05766.
 PF 17-APR-1998; 98US-0061709.
 PR (LUDW-) LUDWIG INST CANCER RES.

XX Chen Y, Gure A, Tsang S, Stockert E, Jager E, Knuth A, Old LJ;
 WPI; 2000-013284/01.

XX Nucleotides representing cancer-associated genes, used to develop
 PT products for the diagnosis, monitoring and treatment of cancers -
 XX Claim 55; Page 42-43; 44pp; English.

CC The present sequence represents an alternative form of a cancer
 CC associated antigen gene designated KOC-3. The specification also
 CC describes a cancer associated antigen designated CT7. The CT7
 CC polynucleotide was isolated from SK-MEL-37 melanoma cells. The
 CC polypeptide has some homology with MAGE-10, limited to about 210 carboxy
 CC terminal amino acids. The amino terminal of the protein has a repetitive
 CC pattern, with repeats rich in serine, proline, glutamine and leucine,
 CC and an almost invariable core of the peptide given in AA43877. The CT7
 CC polypeptide can be processed to peptides which provoke lysis by
 CC cytolytic T cells. The polynucleotides and polypeptides can be used for
 CC treating a cancerous condition and screening for or diagnosing cancerous
 CC conditions. The cancer associated antigens can be used as an immunogenic
 CC or vaccine composition with an adjuvant, e.g. a cytokine, a saponin, or
 CC granulocyte macrophage-colony stimulating factor (GM-CSF).

XX Sequence 3283 BP; 945 A; 833 C; 779 G; 725 T; 1 other;

Query Match 88.2%; Score 19.4; DB 21; Length 3283;
 Best Local Similarity 95.2%; Pred. No. 14;
 Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GCTCTTTGGGGACAGGAGT 21
 |||||
 Db 132 GCTCTTTGGGGACAGGAGT 152

RESULT 10
 AA236152
 ID AA236152 standard; DNA; 3412 BP.
 AC AA236152;
 XX 11-FEB-2000 (first entry)
 DT DNA encoding cancer associated antigen KOC-3.

DE Cancer associated antigen; KOC-3; cancer; vaccine; CT7; ss.
 XX Homo sapiens.

XX WO9954738-A1.
 XX 28-OCT-1999.
 PD 16-MAR-1999; 99WO-US05766.
 PF 17-APR-1998; 98US-0061709.
 PR (LUDW-) LUDWIG INST CANCER RES.

XX Chen Y, Gure A, Tsang S, Stockert E, Jager E, Knuth A, Old LJ;
 WPI; 2000-013284/01.

XX Nucleotides representing cancer-associated genes, used to develop
 PT products for the diagnosis, monitoring and treatment of cancers -
 XX Claim 55; Page 41; 44pp; English.

CC The present sequence represents a cancer associated antigen gene
 CC designated KOC-3. The specification also describes a cancer associated
 CC antigen designated CT7. The CT7 polynucleotide was isolated from
 CC SK-MEL-37 melanoma cells. The polypeptide has some homology with
 CC MAGE-10, limited to about 210 carboxy terminal amino acids. The amino
 CC terminal of the protein has a repetitive pattern, with repeats rich in
 CC serine, proline, glutamine and leucine, and an almost invariable core of
 CC the peptide given in AA43877. The CT7 polypeptide can be processed to
 CC peptides which provoke lysis by cytolytic T cells. The polynucleotides
 CC and polypeptides can be used for treating a cancerous condition and
 CC screening for or diagnosing cancerous conditions. The cancer associated
 CC antigens can be used as an immunogenic or vaccine composition with an
 CC adjuvant, e.g. a cytokine, a saponin, or granulocyte macrophage-colony
 CC stimulating factor (GM-CSF).

Mon Apr 21 10:30:22 2003

XX SQ Sequence 3412 BP; 970 A; 887 C; 804 G; 750 T; 1 other;

Query Match 88.2%; Score 19.4; DB 21; Length 3412;
Best Local Similarity 95.2%; Pred. No. 14;
Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GCTCTTTGGGACAGGAAGT 21
|||||
Db 132 GCTCTTTGGGACAGGAAGT 152

RESULT 11
AAS70982
ID AAS70982 standard; cDNA; 3667 BP.
XX AC AAS70982;
XX DT 13-FEB-2002 (first entry)
XX DE DNA encoding novel human diagnostic protein #6786.
XX KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX OS Homo sapiens.
XX FN WO200175067-A2.
XX PD 11-OCT-2001.
XX PF 30-MAR-2001; 2001WO-US08631.
XX PR 31-MAR-2000; 2000US-0540217.
XX PR 23-AUG-2000; 2000US-0649167.
XX PA (HYSE-) HYSEQ INC.
XX PI Drmanac RT, Liu C, Tang YT;
XX PI WPI; 2001-639362/73.
XX PR P-PSDB; ASG06795.
XX PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity -
XX PS Claim 1; SEQ ID No 6786; 103pp; English.
XX CC The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. AAS64197-AAS94564 represent novel human
CC diagnostic coding sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX SQ Sequence 3667 BP; 1081 A; 882 C; 846 G; 858 T; 0 other;

Query Match 88.2%; Score 19.4; DB 23; Length 3667;
Best Local Similarity 95.2%; Pred. No. 14;
Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GCTCTTTGGGACAGGAAGT 21
|||||
Db 498 GCTCTTTGGGACAGGAAGT 518

RESULT 12
AAS26150
ID AAS26150 standard; cDNA; 3694 BP.
XX AC AAS26150;
XX DT 07-NOV-2001 (first entry)
XX DE Human cDNA encoding a novel secreted protein, Seq ID 329.
XX KW Human; immunosuppressive; antiarthritic; ss; antirheumatic;
KW cystostatic; cardiant; vasotropic; cerebroprotective; nootropic;
KW neuroprotective; antibacterial; virucide; fungicide; ophthalmological;
KW vulnery; secreted protein; rheumatoid arthritis;
KW hyperproliferative disorder; cardiovascular disorder; cardiac arrest;
KW cerebrovascular disorder; cerebral ischaemia; angiogenesis;
KW nervous system disorder; Alzheimer's disease; infection; ocular disorder;
KW corneal infection; wound healing; epithelial cell proliferation;
KW skin ageing; food additive; preservative; antiproliferative.
XX OS Homo sapiens.
XX FN WO200155322-A2.
XX PD 02-AUG-2001.
XX PF 17-JAN-2001; 2001WO-US01341.
XX PR 31-JAN-2000; 2000US-0179065.
XX PR 04-FEB-2000; 2000US-0180628.
XX PR 24-FEB-2000; 2000US-0184664.
XX PR 02-MAR-2000; 2000US-0186350.
XX PR 16-MAR-2000; 2000US-0189874.
XX PR 17-MAR-2000; 2000US-0190076.
XX PR 18-APR-2000; 2000US-0198123.
XX PR 19-MAY-2000; 2000US-0205515.
XX PR 07-JUN-2000; 2000US-0209467.
XX PR 28-JUN-2000; 2000US-0214886.
XX PR 30-JUN-2000; 2000US-0215135.
XX PR 07-JUL-2000; 2000US-0216647.
XX PR 07-JUL-2000; 2000US-0216880.
XX PR 11-JUL-2000; 2000US-0217487.
XX PR 11-JUL-2000; 2000US-0217496.
XX PR 26-JUL-2000; 2000US-0218290.
XX PR 26-JUL-2000; 2000US-0220963.
XX PR 26-JUL-2000; 2000US-0220964.
XX PR 14-AUG-2000; 2000US-0224518.
XX PR 14-AUG-2000; 2000US-0224519.
XX PR 14-AUG-2000; 2000US-0225213.
XX PR 14-AUG-2000; 2000US-0225214.
XX PR 14-AUG-2000; 2000US-0225266.
XX PR 14-AUG-2000; 2000US-0225267.
XX PR 14-AUG-2000; 2000US-0225268.
XX PR 14-AUG-2000; 2000US-0225270.
XX PR 14-AUG-2000; 2000US-0225447.
XX PR 14-AUG-2000; 2000US-0225757.
XX PR 14-AUG-2000; 2000US-0225758.
XX PR 14-AUG-2000; 2000US-0225759.
XX PR 18-AUG-2000; 2000US-0226279.
XX PR 22-AUG-2000; 2000US-0226681.
XX PR 22-AUG-2000; 2000US-0226688.
XX PR 22-AUG-2000; 2000US-0227182.
XX PR 23-AUG-2000; 2000US-0227009.

Mon Apr 21 10:30:22 2003

us-09-270-437d-13.rng

DB 155 GCTCTTTGGGACAGGAGCT 175

RESULT 13
AAV32994
ID AAV32994 standard; cDNA; 1212 BP.

XX AC AAV32994;
XX DT
XX DE
XX DE Human RAD54 nucleic acid sequence comprising exon 1-2.
XX DE Human, RAD54; cancer; xeroderma pigmentosum; Bloom syndrome;
XX KW Werner's syndrome; ATR-X; diagnosis; detection; SNF2 superfamily;
XX KW X-linked mental retardation with alpha-thalassemia syndrome; tumour;
XX KW gene therapy; ss.
XX OS Homo sapiens.
XX OS EP844305-A2.
XX FN
XX XX
XX PD 27-MAY-1998.
XX PF 10-NOV-1997; 97EP-0308998.
XX PF 13-NOV-1996; 96US-0030676.
XX PA (SMIK) SMITHKLINE BEECHAM CORP.
XX PA (UVEJE-) UNIV JEFFERSON THOMAS.
XX PI Croce CM, Fishel RA, Rasio D, Robbins DJ;
XX WPI; 1998-274189/25.
XX DR Human hRAD54 DNA and polypeptide - and agonists, antibodies,
XX PT antagonists, etc.
XX PS Claim 1; Page 25; 64pp; English.
XX CC The present sequence represents a specifically claimed partial nucleic
XX CC acid sequence encoding human RAD54 (hRAD54). A method for analysing a
XX CC sample for mutation of DNA encoding hRAD54 has been developed using a
XX CC DNA sequence of at least 15 and no more than 30 consecutive bases of
XX CC the DNA sequence encoding hRAD54. hRAD54 is a gene thought to be present
XX CC in tumours that display allelic imbalance at 1p32, the chromosomal band
XX CC identified as one of four minimal regions of chromosome 1 deletion in
XX CC breast carcinomas. hRAD54 is useful for production of proteins, inter
XX CC alia, that have been identified as novel hRAD54 by homology between the
XX CC amino acid sequence given in AA62186 and known amino acid sequences
XX CC such as yeast RAD54. hRAD54 proteins are used in the treatment of
XX CC cancer, including Xeroderma Pigmentosum and Bloom syndrome, Werner's
XX CC syndrome and X-linked mental retardation with alpha-thalassaemia
XX CC syndrome and breast cancer. hRAD54 polynucleotides are also useful for
XX CC detecting complementary nucleotides for use as a diagnostic agent,
XX CC especially useful for diagnosis of disease or susceptibility to diseases.
XX CC hRAD54 polynucleotide, proteins, agonists and antagonists which are
XX CC proteins are useful in gene therapy.
XX SQ Sequence 1212 BP; 241 A; 360 C; 287 G; 322 T; 2 other;

Query Match 80.9%; Score 17.8; DB 19; Length 1212;
Best Local Similarity 90.5%; Pred. No. 71;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GCTCTTTGGGACAGGAGCT 21
Db 40 GCGCTTTGGGACAGGAGCT 60

RESULT 14
AAZ23905/c
ID AAZ23905 standard; DNA; 11849 BP.

XX AC AAZ23905;
XX DT 25-JAN-2000 (first entry)
XX DE Human LOBO homologue genomic DNA fragment 7.
XX DE LOBO; long bones; bone development; bone extension; skull; osteopathic;
XX KW diagnostic; pharmaceutical; gene therapy; transgenic animal; disease;
XX KW spondyloepiphyseal dysplasia; achondroplasia; human; ds.
XX OS Homo sapiens.
XX OS WO9950284-A2.
XX PD 07-OCT-1999.
XX PF 26-MAR-1999; 99WO-EP02055.
XX PR 27-MAR-1998; 98DE-1013799.
XX (ROSE/) ROSENTHAL A.
XX PA Rosenthal A, Rump A, Hess J, Aigner T, Wirth T;
XX PI WPI; 1999-601320/51.
XX DR Nucleic acids encoding proteins which influence bone development,
XX PT useful for treating and studying bone disorders -
XX PS Example 3; Page 384-391; 391pp; German.
XX CC This invention describes novel nucleic acids (I; designated LOBO (long
XX CC bones)) encoding proteins influencing bone development in mammals. The
XX CC proteins of the invention reduce and/or inactivate bone extension (i.e.
XX CC development), with exception of the skull and have osteopathic activity.
XX CC The nucleic acid molecules, proteins and antibodies can be used in
XX CC diagnostic or pharmaceutical compounds e.g. for gene therapy. The methods
XX CC and nucleic acid molecules, etc. are useful for production of transgenic
XX CC animals, especially a transgenic mouse for the study of diseases
XX CC associated with bone development, e.g. spondyloepiphyseal dysplasia and
XX CC achondroplasia. This sequence encodes a human LOBO protein described
XX CC in the method of the invention.
XX SQ Sequence 11849 BP; 2693 A; 3311 C; 3343 G; 2502 T; 0 other;

Query Match 80.9%; Score 17.8; DB 20; Length 11849;
Best Local Similarity 90.5%; Pred. No. 93;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GCTCTTTGGGACAGGAGCT 21
Db 7320 GCTCTTTGGGACAGGAGCT 7300

RESULT 15
AAH34291/c
ID AAH34291 standard; cDNA; 3036 BP.
XX AC AAH34291;
XX DT 03-SEP-2001 (first entry)
XX DE Human colon cancer antigen encoding cDNA SEQ ID NO:1373.
XX KW Human; colon cancer; colon cancer antigen; diagnosis; detection;
XX KW colorectal carcinoma; ss.
XX OS Homo sapiens.
XX OS WO200122920-A2.
XX FN
XX XX 05-APR-2001.

XX 28-SEP-2000; 2000WO-US26524.
XX
XX 29-SEP-1999; 99US-0157137.
PR 03-NOV-1999; 99US-0163280.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
PA
XX
XX
PI Ruben SM, Barash SC, Birse CE, Rosen CA;
XX
XX WPI; 2001-235357/24.
DR P-PSDB; AAG74886.
XX
XX Nucleic acids encoding 4277 human colon cancer-associated polypeptides,
PT useful for preventing, diagnosing and/or treating colorectal cancers -
XX
XX Claim 1; Page 3084-3085; 9803pp; English.
XX
XX AAH32943 to AAH37195 and AAG73514 to AAG77788 represent human colon
CC cancer-associated nucleic acid molecules (N) and proteins (P), where
CC the proteins are collectively known as colon cancer antigens. The colon
CC cancer antigens have cytostatic activity and can be used in gene
CC therapy and vaccine production. N and P may be used in the prevention,
CC diagnosis and treatment of diseases associated with inappropriate P
CC expression. For example, N and P may be used to treat disorders
CC associated with decreased expression by rectifying mutations or deletions
CC in a patient's genome that affect the activity of P by expressing
CC inactive proteins or to supplement the patient's own production of P.
CC Additionally, N may be used to produce the colon cancer-associated P,
CC by inserting the nucleic acids into a host cell and culturing the cell
CC to express the proteins. N and P can be used in the prevention, diagnosis
CC and treatment of colorectal carcinomas and cancers. AAH37196 to AAH37204
CC and AAG77789 represent sequences used in the exemplification of the
CC present invention.
CC N.B. Pages 666 to 682 and page 7053 of the sequence listing were
CC missing at time of publication, meaning no sequences are present for
CC SEQ ID NO:1027 to 1052, 7921 and 7922.
XX
SQ Sequence 3036 BP; 961 A; 603 C; 633 G; 830 T; 9 other;
Query Match 79.1%; Score 17.4; DB 22; Length 3036;
Best Local Similarity 94.7%; Pred. No. 1.2e+02;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1 GCTCTTTGGGACAGGAAG 19
|||
Db 2881 GCCTTTGGGACAGGAAG 2863

Search completed: April 18, 2003, 05:45:32
Job time : 104 secs

GenCore version 5.1.4 p5 4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 18, 2003, 04:53:41 ; Search time 21.5 Seconds
(without alignments)
313.809 Million cell updates/sec

Title: US-09-270-437D-13
Perfect score: 22
Sequence: 1 gctcttggggacaggagtc 22

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA:
1: /cgn2_6/ptodata/1/ina/5A COMB.seq.*
2: /cgn2_6/ptodata/1/ina/5B COMB.seq.*
3: /cgn2_6/ptodata/1/ina/6A COMB.seq.*
4: /cgn2_6/ptodata/1/ina/6B COMB.seq.*
5: /cgn2_6/ptodata/1/ina/PTCUS COMB.seq.*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	19.4	88.2	2935	1	US-08-281-193-1
2	19.4	88.2	2935	1	US-08-422-106-1
3	19.4	88.2	2935	2	US-08-735-716-1
4	19.4	88.2	2935	2	US-08-555-568B-1
5	19.4	88.2	2935	4	US-09-519-223-1
6	19.4	88.2	2935	5	PCT-US95-08069-1
7	19.4	88.2	3283	4	US-09-061-709-8
8	19.4	88.2	3412	4	US-09-473-717-1
9	17.2	78.2	4985	4	US-09-398-193-98
10	17.2	78.2	5515	4	US-08-328-862-2
11	16.4	74.5	927	4	US-08-140-729A-2
12	16.4	74.5	1680	1	US-08-546-666-2
13	16.4	74.5	1680	2	US-08-516-745-2
14	16.4	74.5	1680	2	US-09-042-929-2
15	16.4	74.5	1680	2	US-08-546-661-2
16	16.4	74.5	1680	2	US-09-042-960-2
17	16.4	74.5	1680	3	US-09-198-650-2
18	16.4	74.5	1680	3	US-09-042-913-2
19	16.4	74.5	1680	3	US-09-042-937-2
20	16.4	74.5	1680	4	US-09-042-709A-2
21	16.2	73.6	1277	2	US-08-555-568B-18
22	16.2	73.6	1277	4	US-09-519-223-18
23	16.2	73.6	1707	1	US-08-790-309-1
24	16.2	73.6	1707	4	US-09-250-585A-1
25	16.2	73.6	1710	2	US-08-912-794-1
26	16.2	73.6	2109	2	US-08-555-568B-20
27	16.2	73.6			

28	16.2	73.6	2109	4	US-09-519-223-20	Sequence 20, Appl
29	16.2	73.6	2112	2	US-08-555-568B-22	Sequence 22, Appl
30	16.2	73.6	2112	4	US-09-519-223-22	Sequence 22, Appl
31	16.2	73.6	3240	4	US-08-551-896-10	Sequence 10, Appl
32	16.2	73.6	70000	4	US-09-851-896-3	Sequence 3, Appl
C 33	16	72.7	1641	3	US-08-705-771-9	Sequence 9, Appl
C 34	16	72.7	1662	1	US-08-336-408B-5	Sequence 5, Appl
C 35	16	72.7	1662	5	PCT-US91-00399-5	Sequence 5, Appl
C 36	15.8	71.8	401	4	US-09-221-298-40	Sequence 40, Appl
C 37	15.8	71.8	446	4	US-09-641-638-500	Sequence 500, Appl
C 38	15.8	71.8	459	4	US-09-443-041A-13	Sequence 13, Appl
C 39	15.8	71.8	658	4	US-09-385-982-94	Sequence 94, Appl
C 40	15.8	71.8	1447	4	US-09-443-041A-27	Sequence 27, Appl
C 41	15.8	71.8	2002	2	US-08-715-202A-3	Sequence 3, Appl
C 42	15.8	71.8	2002	4	US-09-328-775-3	Sequence 3, Appl
C 43	15.8	71.8	3770	4	US-09-221-017B-807	Sequence 807, Appl
C 44	15.6	70.9	2349	2	US-08-974-546-2	Sequence 2, Appl
C 45	15.6	70.9	3490	2	US-08-841-483-3	Sequence 3, Appl

ALIGNMENTS

RESULT 1
US-08-281-193-1
; Sequence 1, Application US/08281193
; Patent No. 5466595
; GENERAL INFORMATION:
; APPLICANT: Jones, Simon
; APPLICANT: Tang, Jim
; TITLE OF INVENTION: Calcium Independent Phospholipase A2/B
; NUMBER OF SEQUENCES: 15
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/281,193
; FILING DATE:
; CLASSIFICATION: 435
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2935 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 96...2352
US-08-281-193-1

Query Match 88.2%; Score 19.4; DB 1; Length 2935;
Best Local Similarity 95.2%; Pred. No. 1.3;
Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 2 CTCCTTGGGGACAGGAGTC 22
Db 2006 CTCCTTGGGGACAGGAGTC 2026

RESULT 2
US-08-422-106-1
; Sequence 1, Application US/08422106
; Patent No. 5589170
; GENERAL INFORMATION:
; APPLICANT: Jones, Simon
; APPLICANT: Tang, Jim
; TITLE OF INVENTION: Calcium Independent Phospholipase A2/B
; NUMBER OF SEQUENCES: 15

Mon Apr 21 10:30:23 2003

us-09-270-437d-13.rni

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/422,106
FILING DATE: 14-APR-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/281,193
FILING DATE: 27-JUL-1994
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 2935 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
FEATURE:
NAME/KEY: CDS
LOCATION: 96..2352
US-08-422-106-1

Query Match 88.2%; Score 19.4; DB 1; Length 2935;
Best Local Similarity 95.2%; Pred. No. 1.3;
Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 CTCTTTGGGACAGGAGGTC 22
DB 2006 CTCTCTGGGGACAGGAGGTC 2026

RESULT 3
US-08-735-716-1
Sequence 1, Application US/08735716
Patent No. 5840511
GENERAL INFORMATION:
APPLICANT: Jones, Simon
APPLICANT: Tang, Jim
TITLE OF INVENTION: Calcium Independent Phospholipase A2/B
NUMBER OF SEQUENCES: 15
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/735,716
FILING DATE: 23-OCT-1996
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/281,193
FILING DATE: 27-JUL-1994
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 2935 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
FEATURE:
NAME/KEY: CDS
LOCATION: 96..2352
US-08-735-716-1

Query Match 88.2%; Score 19.4; DB 2; Length 2935;
Best Local Similarity 95.2%; Pred. No. 1.3;
Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 CTCTTTGGGACAGGAGGTC 22
DB 2006 CTCTCTGGGGACAGGAGGTC 2026

RESULT 4
US-08-555-568B-1
Sequence 1, Application US/08555568B
Patent No. 5976854
GENERAL INFORMATION:
APPLICANT: Jones, Simon
APPLICANT: Tang, Jim
TITLE OF INVENTION: Calcium Independent Phospholipase A2/B
NUMBER OF SEQUENCES: 25
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genetics Institute, Inc.
STREET: 87 CambridgePark Drive
CITY: Cambridge
STATE: Massachusetts
COUNTRY: U.S.A.
ZIP: 02140
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/555,568B
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Brown, Scott A.
REGISTRATION NUMBER: 32,724
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 498-8224
TELEFAX: (617) 876-5851
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 2935 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
FEATURE:
NAME/KEY: CDS
LOCATION: 96..2352
US-08-555-568B-1

Query Match 88.2%; Score 19.4; DB 2; Length 2935;
Best Local Similarity 95.2%; Pred. No. 1.3;
Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 CTCTTTGGGACAGGAGGTC 22
DB 2006 CTCTCTGGGGACAGGAGGTC 2026

RESULT 5
US-09-519-223-1
Sequence 1, Application US/09519223
Patent No. 6274140
GENERAL INFORMATION:
APPLICANT: Jones, Simon
APPLICANT: Tang, Jim
TITLE OF INVENTION: Calcium Independent Phospholipase A2/B
NUMBER OF SEQUENCES: 25
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genetics Institute, Inc.
STREET: 87 CambridgePark Drive
CITY: Cambridge

Query Match 88.2%; Score 19.4; DB 2; Length 2935;
Best Local Similarity 95.2%; Pred. No. 1.3;
Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;


```

; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02140
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/519,223
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/555,568
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Brown, Scott A.
; REGISTRATION NUMBER: 32,724
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 498-8224
; TELEFAX: (617) 876-5851
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2935 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 96..2352
; US-09-519-223-1

Query Match      88.2%; Score 19.4; DB 4; Length 2935;
Best Local Similarity 95.2%; Pred. No. 1.3;
Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 CTCCTTTGGGACAGGAAGTC 22
Db 2006 CTCCTTTGGGACAGGAAGTC 2026

RESULT 6
PCT-US95-08069-1
; Sequence 1, Application PC/TUS9508069
; GENERAL INFORMATION:
; APPLICANT: GENETICS INSTITUTE, INC.
; TITLE OF INVENTION: Calcium Independent Phospholipase A2/B
; NUMBER OF SEQUENCES: 15
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/08069
; FILING DATE:
; CLASSIFICATION:
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2935 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 96..2352
; PCT-US95-08069-1

Query Match      88.2%; Score 19.4; DB 4; Length 2935;
Best Local Similarity 95.2%; Pred. No. 1.3;
Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 CTCCTTTGGGACAGGAAGTC 22
Db 2006 CTCCTTTGGGACAGGAAGTC 2026

RESULT 7
US-09-061-709-8
; Sequence 8, Application US/09061709B
; Patent No. 6297364
; GENERAL INFORMATION:
; APPLICANT: Chen, Yao-Tsang
; APPLICANT: Gure, Ali
; APPLICANT: Tsang, Solam
; APPLICANT: Stockert, Elisabeth
; APPLICANT: Jager, Elke
; APPLICANT: Knuth, Alexander
; APPLICANT: Old, Lloyd J.
; TITLE OF INVENTION: Isolated Nucleic Acid Molecules Encoding Cancer Associated
; FILE REFERENCE: LUD 5538
; CURRENT APPLICATION NUMBER: US/09/061,709B
; CURRENT FILING DATE: 1998-04-17
; NUMBER OF SEQ ID NOS: 8
; SEQ ID NO 8
; LENGTH: 3283
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; US-09-061-709-8

Query Match      88.2%; Score 19.4; DB 4; Length 3283;
Best Local Similarity 95.2%; Pred. No. 1.4;
Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GCTCTTTGGGACAGGAAGT 21
Db 132 GCTCTTTGGGACAGGAAGT 152

RESULT 8
US-09-061-709-6
; Sequence 6, Application US/09061709B
; Patent No. 6297364
; GENERAL INFORMATION:
; APPLICANT: Chen, Yao-Tsang
; APPLICANT: Gure, Ali
; APPLICANT: Tsang, Solam
; APPLICANT: Stockert, Elisabeth
; APPLICANT: Jager, Elke
; APPLICANT: Knuth, Alexander
; APPLICANT: Old, Lloyd J.
; TITLE OF INVENTION: Isolated Nucleic Acid Molecules Encoding Cancer Associated
; FILE REFERENCE: LUD 5538
; CURRENT APPLICATION NUMBER: US/09/061,709B
; CURRENT FILING DATE: 1998-04-17
; NUMBER OF SEQ ID NOS: 8
; SEQ ID NO 6
; LENGTH: 3412
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; US-09-061-709-6

Query Match      88.2%; Score 19.4; DB 4; Length 3412;
Best Local Similarity 95.2%; Pred. No. 1.4;
Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GCTCTTTGGGACAGGAAGT 21
Db 132 GCTCTTTGGGACAGGAAGT 152
```

us-09-270-437d-13.rni

Mon Apr 21 10:30:23 2003

Db 4273 GCACCTTTGGGACAGGAGGTC 4252

RESULT 11

US-08-928-862-2/c

; Sequence 2, Application US/08928862

; Patent No. 6309877

; GENERAL INFORMATION:

; APPLICANT: Chau, Raymond M. W.

; TITLE OF INVENTION: Isolation and Use of Motoneuronotrophic Factors

; FILE REFERENCE: 12592-2

; CURRENT APPLICATION NUMBER: US/08/928,862

; CURRENT FILING DATE: 1997-09-12

; NUMBER OF SEQ ID NOS: 5

; SOFTWARE: Patentin Ver. 2.0

; SEQ ID NO 2

; LENGTH: 927

; TYPE: DNA

; ORGANISM: Homo sapiens

US-08-928-862-2

Query Match 74.5%; Score 16.4; DB 4; Length 927;

Best Local Similarity 94.4%; Pred. No. 32;

Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 CTTTGGGACAGGAGGTC 21

Db 200 CTTTGGGACAGGAGGTC 183

RESULT 12

US-08-140-729A-2/c

; Sequence 2, Application US/08140729A

; Patent No. 5658782

; GENERAL INFORMATION:

; APPLICANT: Amara, Susan G

; APPLICANT: Arriza, Jeffrey L

; TITLE OF INVENTION: Amino Acid Transporters and Uses

; NUMBER OF SEQUENCES: 17

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Allegretti & Witcoff, Ltd.

; STREET: 10 South Wacker Drive, Suite 3000

; CITY: Chicago

; STATE: IL USA

; COUNTRY: USA

; ZIP: 60606

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patentin Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/140,729A

; FILING DATE: 20 OCT 1993

; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:

; NAME: No. 5658782han, Kevin E

; REGISTRATION NUMBER: 35,303

; REFERENCE/DOCKET NUMBER: 93,509

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 312-715-1000

; TELEFAX: 312-715-1234

; TELEX: 910-221-5317

; INFORMATION FOR SEQ ID NO: 2:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 1680 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: cDNA

; FEATURE:

; NAME/KEY: 5'UTR

; LOCATION: 1..30

QY 1 GCTCTTTGGGACAGGAGGTC 21

Db 132 GCTCTTTGGGACAGGAGGTC 152

RESULT 9

US-09-473-717-1/c

; Sequence 1, Application US/09473717

; Patent No. 6372475

; GENERAL INFORMATION:

; APPLICANT: Storm, Daniel R.

; APPLICANT: Hacker, Beth

; APPLICANT: Tomlinson, James E.

; APPLICANT: University of Washington

; TITLE OF INVENTION: CLONING AND CHARACTERIZATION OF A HUMAN ADENYLYL

; FILE REFERENCE: 44481-5029-01-US

; CURRENT APPLICATION NUMBER: US/09/473,717

; CURRENT FILING DATE: 1999-12-29

; PRIOR FILING DATE: 1998-07-01

; PRIOR APPLICATION NUMBER: 60/098,559

; PRIOR FILING DATE: 1997-07-01

; PRIOR APPLICATION NUMBER: 08/886,440

; PRIOR FILING DATE: 1997-07-01

; NUMBER OF SEQ ID NOS: 3

; SOFTWARE: Patentin Ver. 2.0

; SEQ ID NO 1

; LENGTH: 4985

; TYPE: DNA

; ORGANISM: human type IX adenylyl cyclase

; FEATURE:

; NAME/KEY: CDS

; LOCATION: (17) .. (3898)

US-09-473-717-1

Query Match 78.2%; Score 17.2; DB 4; Length 4985;

Best Local Similarity 86.4%; Pred. No. 18;

Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 GCTCTTTGGGACAGGAGGTC 22

Db 3751 GCACCTTTGGGACAGGAGGTC 3730

RESULT 10

US-09-398-193-98/c

; Sequence 98, Application US/09398193

; Patent No. 6197581

; GENERAL INFORMATION:

; APPLICANT: Medical Research Council

; TITLE OF INVENTION: Adenylyl cyclase and uses therefor

; FILE REFERENCE: P24360-

; CURRENT APPLICATION NUMBER: US/09/398,193

; CURRENT FILING DATE: 1999-09-17

; NUMBER OF SEQ ID NOS: 104

; SOFTWARE: Patentin Ver. 2.1

; SEQ ID NO 98

; LENGTH: 5515

; TYPE: DNA

; ORGANISM: Human

; FEATURE:

; NAME/KEY: CDS

; LOCATION: (539) .. (4600)

US-09-398-193-98

Query Match 78.2%; Score 17.2; DB 4; Length 5515;

Best Local Similarity 86.4%; Pred. No. 18;

Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 GCTCTTTGGGACAGGAGGTC 22

us-09-270-437d-13.rnd

Mon Apr 21 10:30:23 2003

```

; Patent No. 5919628
; GENERAL INFORMATION:
; APPLICANT: Amara, Susan G
; TITLE OF INVENTION: Amino Acid Transporters and Uses
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Allegretti & Witcoff, Ltd.
; STREET: 10 South Wacker Drive, Suite 3000
; CITY: Chicago
; STATE: IL
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/042,929
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/140,729
; FILING DATE: 20 OCT 1993
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 5919628nan, Kevin E
; REGISTRATION NUMBER: 35,303
; REFERENCE/DOCKET NUMBER: 93,509
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-715-1000
; TELEFAX: 312-715-1234
; TELEX: 910-221-5317
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1680 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: 5'UTR
; LOCATION: 1..30
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 31..1626
; FEATURE:
; NAME/KEY: 3'UTR
; LOCATION: 1626..1680
; US-09-042-929-2

```

```

Query Match 74.5%; Score 16.4; DB 2; Length 1680;
Best Local Similarity 94.4%; Pred. No. 36;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

QY 2 CTCCTTGGGGACAGGAG 19
Db 528 CTCCTTGGGGACAGGAG 511

```

```

Search completed: April 18, 2003, 07:34:33
Job time : 24.5 secs

```

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 18, 2003, 06:35:31 ; Search time 91.3333 Seconds
(without alignments)
241.975 Million cell updates/sec

Title: US-09-270-437D-13

Perfect score: 22

Sequence: 1 gctcttggggacaggagtc 22

Scoring table: IDENTITY NUC

Gapop 10_0 , Gapext 1.0

Searched: 639749 seqs, 502280978 residues

Total number of hits satisfying chosen parameters: 1279498

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications NA:*

- 1: /cgn2_6/ptodata/1/pubna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/1/pubna/PCT_NEW_PUB.seq:*
- 3: /cgn2_6/ptodata/1/pubna/US06_NEW_PUB.seq:*
- 4: /cgn2_6/ptodata/1/pubna/US06_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/1/pubna/US07_NEW_PUB.seq:*
- 6: /cgn2_6/ptodata/1/pubna/PCTUS_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/1/pubna/US08_NEW_PUB.seq:*
- 8: /cgn2_6/ptodata/1/pubna/US08_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/1/pubna/US09_NEW_PUB.seq:*
- 10: /cgn2_6/ptodata/1/pubna/US09_PUBCOMB.seq:*
- 11: /cgn2_6/ptodata/1/pubna/US10_NEW_PUB.seq:*
- 12: /cgn2_6/ptodata/1/pubna/US10_PUBCOMB.seq:*
- 13: /cgn2_6/ptodata/1/pubna/US60_NEW_PUB.seq:*
- 14: /cgn2_6/ptodata/1/pubna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	19.4	88.2	2935	10 US-09-927-180-1	Sequence 1, Appli
2	19.4	88.2	3283	10 US-09-899-651-8	Sequence 8, Appli
3	19.4	88.2	3412	10 US-09-899-651-6	Sequence 6, Appli
4	19.4	88.2	3694	10 US-09-764-864-329	Sequence 329, App
5	17.2	78.2	4985	12 US-10-071-223-1	Sequence 1, Appli
6	17.2	78.2	5515	10 US-09-751-100B-98	Sequence 98, Appli
7	16.8	76.4	401	9 US-09-946-807-702	Sequence 702, App
8	16.8	76.4	401	10 US-09-795-668-702	Sequence 702, App
9	16.8	76.4	401	10 US-09-795-686-702	Sequence 702, App
10	16.8	76.4	80959	9 US-09-858-546-3	Sequence 3, Appli
11	16.4	74.5	15531	9 US-10-125-540-600	Sequence 600, App
12	16.4	74.5	15531	9 US-09-764-870-600	Sequence 600, App
13	16.4	74.5	659158	9 US-09-771-208-20	Sequence 20, Appl
14	16.2	73.6	256	10 US-09-878-574-5931	Sequence 5931, Ap
15	16.2	73.6	275	10 US-09-878-574-9673	Sequence 9673, Ap
16	16.2	73.6	694	10 US-09-878-574-4815	Sequence 4815, Ap
17	16.2	73.6	1277	10 US-09-927-180-18	Sequence 18, Appl
18	16.2	73.6	2109	10 US-09-927-180-20	Sequence 20, Appl
19	16.2	73.6	2112	10 US-09-927-180-22	Sequence 22, Appl

20	16.2	73.6	3231	10 US-09-864-864-306	Sequence 306, App
21	16.2	73.6	24707	9 US-10-274-968-3	Sequence 3, Appli
22	16.2	73.6	24707	10 US-09-740-027-3	Sequence 3, Appli
23	16	72.7	361	9 US-09-736-457-1159	Sequence 1159, Ap
24	16	72.7	361	9 US-09-902-941-1159	Sequence 1159, Ap
25	16	72.7	361	9 US-09-849-626-1159	Sequence 1159, Ap
26	16	72.7	361	9 US-10-017-754-1159	Sequence 1159, Ap
27	16	72.7	437	10 US-09-917-800A-392	Sequence 392, App
28	16	72.7	669	10 US-09-917-800A-486	Sequence 486, App
29	16	72.7	1594	10 US-09-964-824A-559	Sequence 559, App
30	15.8	71.8	339	9 US-09-803-719-978	Sequence 978, App
31	15.8	71.8	401	9 US-10-025-380-40	Sequence 40, Appl
32	15.8	71.8	401	10 US-09-922-217-40	Sequence 40, Appl
33	15.8	71.8	401	10 US-09-833-263-40	Sequence 40, Appl
34	15.8	71.8	409	10 US-09-960-352-9154	Sequence 9154, Ap
35	15.8	71.8	412	10 US-09-960-352-14119	Sequence 14119, A
36	15.8	71.8	449	10 US-09-998-598-2037	Sequence 2037, Ap
37	15.8	71.8	527	10 US-09-920-300A-1339	Sequence 1339, Ap
38	15.8	71.8	527	12 US-10-033-528-1339	Sequence 1339, Ap
39	15.8	71.8	689	10 US-09-917-800A-1159	Sequence 1159, Ap
40	15.8	71.8	787	9 US-10-025-380-132	Sequence 132, App
41	15.8	71.8	787	10 US-09-922-217-132	Sequence 132, App
42	15.8	71.8	787	10 US-09-833-263-132	Sequence 132, App
43	15.8	71.8	1369	10 US-09-822-849A-67	Sequence 67, Appl
44	15.8	71.8	1787	10 US-09-925-302-3	Sequence 3, Appli
45	15.8	71.8	1923	10 US-09-822-849A-506	Sequence 506, App

ALIGNMENTS

RESULT 1

US-09-927-180-1
; Sequence 1, Application US/09927180
; Patent No. US20020106364A1
; GENERAL INFORMATION:
; APPLICANT: Jones, Simon
; Tang, Jim
; TITLE OF INVENTION: Calcium Independent Phospholipase A2/B
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genetics Institute, Inc.
; STREET: 87 CambridgePark Drive
; CITY: Cambridge
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02140
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (BPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/927,180
; FILING DATE: 09-Aug-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/519,223
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Brown, Scott A.
; REGISTRATION NUMBER: 32,724
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 498-8224
; TELEFAX: (617) 876-5851
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2935 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO

us-09-270-437d-13.rnpb

Mon Apr 21 10:30:25 2003

```
; ANTI-SENSE: NO
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 96..2352
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-927-180-1

Query Match      88.2%; Score 19.4; DB 10; Length 2935;
Best Local Similarity 95.2%; Pred. No. 4.3;
Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 CTCCTTTGGGACAGGAGGTC 22
    |||||
Db 2006 CTCCTTTGGGACAGGAGGTC 2026

RESULT 2
US-09-899-651-8
; Sequence 8, Application US/09899651
; Patent No. US20020111470A1
; GENERAL INFORMATION:
; APPLICANT: Chen, Yao-Tseng
; APPLICANT: Gure, Ali
; APPLICANT: Tsang, Solam
; APPLICANT: Stockert, Elisabeth
; APPLICANT: Jager, Elke
; APPLICANT: Knuth, Alexander
; APPLICANT: Old, Lloyd J.
; TITLE OF INVENTION: Isolated Nucleic Acid Molecules Encoding Cancer
; TITLE OF INVENTION: Associated
; TITLE OF INVENTION: Antigen, The Antigens Per Se, And Uses Thereof
; FILE REFERENCE: LUD 5538
; CURRENT APPLICATION NUMBER: US/09/899,651
; CURRENT FILING DATE: 2001-07-06
; PRIOR APPLICATION NUMBER: US/09/061,709
; PRIOR FILING DATE: 1998-04-17
; NUMBER OF SEQ ID NOS: 8
; SEQ ID NO 8
; LENGTH: 3283
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
US-09-899-651-8

Query Match      88.2%; Score 19.4; DB 10; Length 3283;
Best Local Similarity 95.2%; Pred. No. 4.3;
Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GCTCTTTGGGACAGGAGGT 21
    |||||
Db 132 GCTCTTTGGGACAGGAGGT 152

RESULT 3
US-09-899-651-6
; Sequence 6, Application US/09899651
; Patent No. US20020111470A1
; GENERAL INFORMATION:
; APPLICANT: Chen, Yao-Tseng
; APPLICANT: Gure, Ali
; APPLICANT: Tsang, Solam
; APPLICANT: Stockert, Elisabeth
; APPLICANT: Jager, Elke
; APPLICANT: Knuth, Alexander
; APPLICANT: Old, Lloyd J.
; TITLE OF INVENTION: Isolated Nucleic Acid Molecules Encoding Cancer
; TITLE OF INVENTION: Associated
; TITLE OF INVENTION: Antigen, The Antigens Per Se, And Uses Thereof
; FILE REFERENCE: LUD 5538
; CURRENT APPLICATION NUMBER: US/09/899,651
; CURRENT FILING DATE: 2001-07-06
; PRIOR APPLICATION NUMBER: US/09/061,709
; PRIOR FILING DATE: 1998-04-17
; NUMBER OF SEQ ID NOS: 8
; SEQ ID NO 8
; LENGTH: 3283
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
US-09-899-651-6

Query Match      88.2%; Score 19.4; DB 10; Length 3283;
Best Local Similarity 95.2%; Pred. No. 4.3;
Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GCTCTTTGGGACAGGAGGT 21
    |||||
Db 132 GCTCTTTGGGACAGGAGGT 152

RESULT 4
US-09-764-864-329
; Sequence 329, Application US/09764864
; Patent No. US20020132753A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PTZ23
; CURRENT APPLICATION NUMBER: US/09/764,864
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1792
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 329
; LENGTH: 3694
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
US-09-764-864-329

Query Match      88.2%; Score 19.4; DB 10; Length 3694;
Best Local Similarity 95.2%; Pred. No. 4.3;
Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GCTCTTTGGGACAGGAGGT 21
    |||||
Db 132 GCTCTTTGGGACAGGAGGT 152

RESULT 5
US-10-071-223-1/c
; Sequence 1, Application US/10071223
; Patent No. US20020137174A1
; GENERAL INFORMATION:
; APPLICANT: Storm, Daniel R.
; APPLICANT: Hacker, Beth
; APPLICANT: Tomlinson, James E.
; APPLICANT: COR Therapeutics, Inc.
; APPLICANT: University of Washington
; TITLE OF INVENTION: CLONING AND CHARACTERIZATION OF A HUMAN ADENYLYL
; FILE REFERENCE: 44481-5029-02-US
; CURRENT APPLICATION NUMBER: US/10/071,223
; CURRENT FILING DATE: 2002-02-11
; PRIOR APPLICATION NUMBER: 09/473,717
; PRIOR FILING DATE: 1999-12-29
; PRIOR APPLICATION NUMBER: PCT/US98/13541
; PRIOR FILING DATE: 1998-07-01
; PRIOR APPLICATION NUMBER: 60/098,559
; PRIOR FILING DATE: 1997-07-01
; PRIOR APPLICATION NUMBER: 08/886,440
; PRIOR FILING DATE: 1997-07-01
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 4985
; TYPE: DNA
; FEATURE:
US-10-071-223-1/c
```

```

; ORGANISM: human type IX adenylyl cyclase
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (17)..(3898)
US-10-071-223-1

```

```
Query Match      78.2%; Score 17.2; DB 12; Length 4985;
Best Local Similarity 86.4%; Pred. NO. 44;
Matches 19; Conservative 0; Mismatches 3; Indels 0;
```

QY 1 GCTCTTTGGGACAGGAAGGTC 22
||| ||| ||| ||| ||| ||| ||| |||
Db 3751 GCACCTTTGGGTACAGGTAGGTC 3730

RESULT 6

US-09-751-100B-98/c
; Sequence 98, Application US/09751100B
; Patent No. US20020142436A1

GENERAL INFORMATION:

```

; APPLICANT: Medical Research Council
; TITLE OF INVENTION: Human Adenylate Cyclase and Use Therefor
; FILE REFERENCE: P27948A
; CURRENT APPLICATION NUMBER: US/09/751.100B
; CURRENT FILING DATE: 2000-12-28
; NUMBER OF SEQ ID NOS: 104
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 98
; LENGTH: 5515

```

TYPE: DNA

```

ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (539)..(4600)
OTHER INFORMATION:
US-09-751-100B-98

```

```

Query Match          78.2%; Score 17.2; DB 10; Length 5515;
Best Local Similarity 86.4%; Pred. No. 44;
Matches 19; Conservative 0; Mismatches 3; Indels 0;

```

1 GCTCTTTGGGACAGGAAGGTC 22
 4273 GCACCTTTGGGTACAGGTAGGTC 4252

RESULT 7

US-09-946-807-702/c
 Sequence 702, Application US/09946807
 Patent No. US20020165144A1
 GENERAL INFORMATION:
 APPLICANT: Stefansson, Hreinn
 APPLICANT: Steinthorsdottir, Valgerdur
 APPLICANT: Gulcher, Jeffrey R.
 TITLE OF INVENTION: HUMAN SCHIZOPHRENIA GENE
 FILE REFERENCE: 2345.2004-001
 CURRENT APPLICATION NUMBER: US/09/946,807
 CURRENT FILING DATE: 2001-09-05
 PRIOR APPLICATION NUMBER: US/09/795,668
 PRIOR FILING DATE: 2001-02-28
 PRIOR APPLICATION NUMBER: US/09/515,716
 PRIOR FILING DATE: 2000-02-28
 NUMBER OF SEQ ID NOS: 1531
 SOFTWARE: FastSeq for Windows Version 4.0
 SEQ ID NO 702
 LENGTH: 401
 TYPE: DNA
 ORGANISM: Homo sapiens
 3-09-946-807-702

Query Match 76.4%; Score 15.8; DB 9; Length 401;
Best Local Similarity 90.0%;
Matches 18; Conservative 0; Mismatches 2; Indels

Qy 3 TCTTTGGGGACAGGAAGTC 22
| | | | | | | | | |
Db 98 TGTTTAGGGACAGGAAGTC 79

RESULT 8

```

US-09-795-668-702/c
; Sequence 702, Application US/09795668
; Patent No. US20020045577A1
; GENERAL INFORMATION:
; APPLICANT: Stefansson, Hreinn.
; APPLICANT: Steinhorsdottir, Valgerdur
; APPLICANT: Gulcher, Jeffrey R.
; TITLE OF INVENTION: HUMAN SCHIZOPHRENIA GENE
; FILE REFERENCE: 2345.2004-001
; CURRENT APPLICATION NUMBER: US/09/795.668
; CURRENT FILING DATE: 2001-02-28
; PRIOR FILING DATE: 2000-02-28
; NUMBER OF SEQ ID NOS: 1531
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 702
; LENGTH: 401
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-795-668-702

```

```
Query Match 76.4%; Score 16.8; DB 10; Length 401;
Best Local Similarity 90.0%; Pred. No. 75;
Matches 18; Conservative 0; Mismatches 2; Indels 0
```

QY	3	TCTTTGGGACAGGAAGTTC	22
Db	98	TGTTTAGGACAGGAAGTTC	79

RESULT 9

US-09-795-686-702/c
Sequence 702, Application US/09795686
Patent No. US20020094954A1
GENERAL INFORMATION:
APPLICANT: Stefansson, Hreinn
APPLICANT: Steinthorsdottir, Valgerdur
APPLICANT: Gulcher, Jeffrey R.
TITLE OF INVENTION: HUMAN SCHIZOPHRENIA GENE
FILE REFERENCE: 2345 2005-001
CURRENT APPLICATION NUMBER: US/09/795,686
CURRENT FILING DATE: 2001-02-28
PRIOR APPLICATION NUMBER: US 09/515,715
PRIOR FILING DATE: 2000-02-28
NUMBER OF SEQ ID NOS: 1531
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 702
LENGTH: 401
TYPE: DNA
ORGANISM: Homo sapiens
S-09-795-686-702

Query Match 76.4%; Score 16.8; DB 10; Length 401;
Best Local Similarity 90.0%; Pred. No. 75;
Matches 18; Conservative 0; Mismatches 2; Indels 0

Y

3 TCTTTGGGACAGGAAGTC 22
| | | | |
b 98 TGTTTAGGGACAGGAAGTC 79

RESULT 10

S-09-858-546-3/c
Sequence 3, Application US/09858546
Patent No. US20020172995A1
GENERAL INFORMATION:

us-09-270-437d-13.rnpb

Mon Apr 21 10:30:25 2003

APPLICANT: SHAO, Wei et al
 TITLE OF INVENTION: ISOLATED HUMAN SECRETED PROTEINS, AND
 TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN SECRETED PROTEINS, AND
 TITLE OF INVENTION: USES THEREOF
 FILE REFERENCE: CL001232
 CURRENT APPLICATION NUMBER: US/09/858,546
 CURRENT FILING DATE: 2001-05-17
 NUMBER OF SEQ ID NOS: 5
 SOFTWARE: FastSeq for Windows Version 4.0
 SEQ ID NO 3
 LENGTH: 80959
 TYPE: DNA
 ORGANISM: Human
 FEATURE:
 NAME/KEY: misc_feature
 LOCATION: (1)...(80959)
 OTHER INFORMATION: n = A,T,C or G
 US-09-858-546-3

Query Match 76.4%; Score 16.8; DB 9; Length 80959;
 Best Local Similarity 90.0%; Pred. No. 61;
 Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GCTCTTTGGGACAGGAGG 20
 |||||
 Db 75694 GCTCTTTGGGACAGGAGG 75675

RESULT 11
 US-10-125-540-600
 Sequence 600, Application US/10125540
 Publication No. US20030059875A1
 GENERAL INFORMATION:
 APPLICANT: Rosen et al.
 TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
 FILE REFERENCE: PT214CI
 CURRENT APPLICATION NUMBER: US/10/125,540
 CURRENT FILING DATE: 2002-04-19
 Prior Application removed - See File Wrapper or Palm
 NUMBER OF SEQ ID NOS: 646
 SOFTWARE: PatentIn Ver. 2.0
 SEQ ID NO 600
 LENGTH: 15531
 TYPE: DNA
 ORGANISM: Homo sapiens
 US-10-125-540-600

Query Match 74.5%; Score 16.4; DB 9; Length 15531;
 Best Local Similarity 94.4%; Pred. No. 99;
 Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 TCTTTGGGACAGGAGG 20
 |||||
 Db 12698 TCTTTGGGACAGGAGG 12715

RESULT 12
 US-09-764-870-600
 Sequence 600, Application US/09764870
 Patent No. US20020042386A1
 GENERAL INFORMATION:
 APPLICANT: Rosen et al.
 TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
 FILE REFERENCE: PT214
 CURRENT APPLICATION NUMBER: US/09/764,870
 CURRENT FILING DATE: 2001-01-17
 Prior application data removed - consult PALM or file wrapper
 NUMBER OF SEQ ID NOS: 646
 SOFTWARE: PatentIn Ver. 2.0
 SEQ ID NO 600
 LENGTH: 15531
 TYPE: DNA
 ORGANISM: Homo sapiens

US-09-764-870-600

Query Match 74.5%; Score 16.4; DB 10; Length 15531;
 Best Local Similarity 94.4%; Pred. No. 99;
 Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 TCTTTGGGACAGGAGG 20
 |||||
 Db 12698 TCTTTGGGACAGGAGG 12715

RESULT 13
 US-09-771-208-20
 Sequence 20, Application US/09771208
 Patent No. US20020155564A1.
 GENERAL INFORMATION:
 APPLICANT: MEDRANO, JUAN
 APPLICANT: BRADFORD, ERIC
 APPLICANT: HORVAT, SIMON
 TITLE OF INVENTION: CLONING OF A HIGH-GROWTH GENE
 FILE REFERENCE: 407T-923710US
 CURRENT APPLICATION NUMBER: US/09/771,208
 CURRENT FILING DATE: 2001-01-26
 PRIOR APPLICATION NUMBER: US 08/999,477
 PRIOR FILING DATE: 1997-12-29
 NUMBER OF SEQ ID NOS: 20
 SOFTWARE: PatentIn version 3.0
 SEQ ID NO 20
 LENGTH: 659158
 TYPE: DNA
 ORGANISM: Mus musculus
 FEATURE:
 NAME/KEY: misc_feature
 LOCATION: (123459)..(123478)
 OTHER INFORMATION: n is unidentified a, c, g, or t
 NAME/KEY: misc_feature
 LOCATION: (602466)..(602485)
 OTHER INFORMATION: n is unidentified a, c, g, or t
 NAME/KEY: misc_feature
 LOCATION: (546998)..(547017)
 OTHER INFORMATION: n is unidentified a, c, g, or t
 NAME/KEY: misc_feature
 LOCATION: (494715)..(494814)
 OTHER INFORMATION: n is unidentified a, c, g, or t
 NAME/KEY: misc_feature
 LOCATION: (390986)..(391005)
 OTHER INFORMATION: n is unidentified a, c, g, or t
 NAME/KEY: misc_feature
 LOCATION: (346860)..(346823)
 OTHER INFORMATION: n is unidentified a, c, g, or t
 NAME/KEY: misc_feature
 LOCATION: (317174)..(317193)
 OTHER INFORMATION: n is unidentified a, c, g, or t
 NAME/KEY: misc_feature
 LOCATION: (280353)..(280373)
 OTHER INFORMATION: n is unidentified a, c, g, or t
 NAME/KEY: misc_feature
 LOCATION: (271829)..(271848)
 OTHER INFORMATION: n is unidentified a, c, g, or t
 NAME/KEY: misc_feature
 LOCATION: (183872)..(183891)
 OTHER INFORMATION: n is unidentified a, c, g, or t
 NAME/KEY: misc_feature
 LOCATION: (170625)..(170645)
 OTHER INFORMATION: n is unidentified a, c, g, or t
 NAME/KEY: misc_feature
 LOCATION: (132680)..(132700)
 OTHER INFORMATION: n is unidentified a, c, g, or t
 NAME/KEY: misc_feature
 OTHER INFORMATION: n is a, c, g, or t
 US-09-771-208-20

Query Match 74.5%; Score 16.4; DB 9; Length 659158;

Best Local Similarity 94.4%; Pred. No. 84;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Job time : 206.5 secs

QY 4 CTTTGGGACAGGAAGGT 21
|||||
Db 39002 CTTTGGGACAGGAAGGT 39019

RESULT 14
US-09-878-574-5931
; Sequence 5931, Application US/09878574
; Patent No. US20020110548A1
; GENERAL INFORMATION:
; APPLICANT: Byrum, Joseph R.
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Thompson, Michael D.
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(15401)B
; CURRENT APPLICATION NUMBER: US/09/878,574
; CURRENT FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: 09/333,535
; PRIOR FILING DATE: 1999-06-14
; NUMBER OF SEQ ID NOS: 15775
; SEQ ID NO 5931
; LENGTH: 256
; TYPE: DNA
; ORGANISM: Glycine max
; OTHER INFORMATION: Clone ID: 701097533H1
US-09-878-574-5931

Query Match 73.6%; Score 16.2; DB 10; Length 256;
Best Local Similarity 85.7%; Pred. No. 1.4e+02;
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 GCTCTTTGGGACAGGAAGGT 21
|||||
Db 145 GCGCCTTGGGACAGGAAGGT 165

RESULT 15
US-09-878-574-9673
; Sequence 9673, Application US/09878574
; Patent No. US20020110548A1
; GENERAL INFORMATION:
; APPLICANT: Byrum, Joseph R.
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Thompson, Michael D.
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(15401)B
; CURRENT APPLICATION NUMBER: US/09/878,574
; CURRENT FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: 09/333,535
; PRIOR FILING DATE: 1999-06-14
; NUMBER OF SEQ ID NOS: 15775
; SEQ ID NO 9673
; LENGTH: 275
; TYPE: DNA
; ORGANISM: Glycine max
; OTHER INFORMATION: Clone ID: 701102760H1
US-09-878-574-9673

Query Match 73.6%; Score 16.2; DB 10; Length 275;
Best Local Similarity 85.7%; Pred. No. 1.4e+02;
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 GCTCTTTGGGACAGGAAGGT 21
|||||
Db 120 GCGCCTTGGGACAGGAAGGT 140

Search completed: April 18, 2003, 10:16:02



GenCore version 5.1.4 p5 4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 18, 2003, 05:48:17 ; Search time 1211 Seconds
(without alignments)
456.759 Million cell updates/sec

Title: US-09-270-437D-13

Perfect score: 22

Sequence: 1 gctcttgggacgaaggc 22

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 24791104 seqs, 12571243825 residues

Total number of hits satisfying chosen parameters: 49582208

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Pending Patents NA Main:

1: /cgn2_6/ptodata/1/pna/PCTUS COMB.seq.*
2: /cgn2_6/ptodata/1/pna/US06 COMB.seq.*
3: /cgn2_6/ptodata/1/pna/US07 COMB.seq.*
4: /cgn2_6/ptodata/1/pna/US08 COMB.seq.*
5: /cgn2_6/ptodata/1/pna/US081 COMB.seq.*
6: /cgn2_6/ptodata/1/pna/US082 COMB.seq.*
7: /cgn2_6/ptodata/1/pna/US083 COMB.seq.*
8: /cgn2_6/ptodata/1/pna/US084 COMB.seq.*
9: /cgn2_6/ptodata/1/pna/US085 COMB.seq.*
10: /cgn2_6/ptodata/1/pna/US086 COMB.seq.*
11: /cgn2_6/ptodata/1/pna/US087 COMB.seq.*
12: /cgn2_6/ptodata/1/pna/US088 COMB.seq.*
13: /cgn2_6/ptodata/1/pna/US089 COMB.seq.*
14: /cgn2_6/ptodata/1/pna/US090 COMB.seq.*
15: /cgn2_6/ptodata/1/pna/US091 COMB.seq.*
16: /cgn2_6/ptodata/1/pna/US092 COMB.seq.*
17: /cgn2_6/ptodata/1/pna/US093 COMB.seq.*
18: /cgn2_6/ptodata/1/pna/US094 COMB.seq.*
19: /cgn2_6/ptodata/1/pna/US095A COMB.seq.*
20: /cgn2_6/ptodata/1/pna/US095B COMB.seq.*
21: /cgn2_6/ptodata/1/pna/US095C COMB.seq.*
22: /cgn2_6/ptodata/1/pna/US095D COMB.seq.*
23: /cgn2_6/ptodata/1/pna/US096A COMB.seq.*
24: /cgn2_6/ptodata/1/pna/US096B COMB.seq.*
25: /cgn2_6/ptodata/1/pna/US096C COMB.seq.*
26: /cgn2_6/ptodata/1/pna/US096D COMB.seq.*
27: /cgn2_6/ptodata/1/pna/US096E COMB.seq.*
28: /cgn2_6/ptodata/1/pna/US096F COMB.seq.*
29: /cgn2_6/ptodata/1/pna/US097A COMB.seq.*
30: /cgn2_6/ptodata/1/pna/US097B COMB.seq.*
31: /cgn2_6/ptodata/1/pna/US097C COMB.seq.*
32: /cgn2_6/ptodata/1/pna/US098A COMB.seq.*
33: /cgn2_6/ptodata/1/pna/US098B COMB.seq.*
34: /cgn2_6/ptodata/1/pna/US098C COMB.seq.*
35: /cgn2_6/ptodata/1/pna/US099A COMB.seq.*
36: /cgn2_6/ptodata/1/pna/US099B COMB.seq.*
37: /cgn2_6/ptodata/1/pna/US099C COMB.seq.*
38: /cgn2_6/ptodata/1/pna/US099D COMB.seq.*
39: /cgn2_6/ptodata/1/pna/US100A COMB.seq.*
40: /cgn2_6/ptodata/1/pna/US100B COMB.seq.*
41: /cgn2_6/ptodata/1/pna/US101A COMB.seq.*
42: /cgn2_6/ptodata/1/pna/US101B COMB.seq.*
43: /cgn2_6/ptodata/1/pna/US102A COMB.seq.*
44: /cgn2_6/ptodata/1/pna/US6000 COMB.seq.*
45: /cgn2_6/ptodata/1/pna/US6001 COMB.seq.*
46: /cgn2_6/ptodata/1/pna/US6002 COMB.seq.*
47: /cgn2_6/ptodata/1/pna/US6003 COMB.seq.*
48: /cgn2_6/ptodata/1/pna/US6004 COMB.seq.*
49: /cgn2_6/ptodata/1/pna/US6005 COMB.seq.*
50: /cgn2_6/ptodata/1/pna/US6006 COMB.seq.*
51: /cgn2_6/ptodata/1/pna/US6007 COMB.seq.*
52: /cgn2_6/ptodata/1/pna/US6008 COMB.seq.*
53: /cgn2_6/ptodata/1/pna/US6009 COMB.seq.*
54: /cgn2_6/ptodata/1/pna/US6010 COMB.seq.*
55: /cgn2_6/ptodata/1/pna/US6011 COMB.seq.*
56: /cgn2_6/ptodata/1/pna/US6012 COMB.seq.*
57: /cgn2_6/ptodata/1/pna/US6013 COMB.seq.*
58: /cgn2_6/ptodata/1/pna/US6014 COMB.seq.*
59: /cgn2_6/ptodata/1/pna/US6015 COMB.seq.*
60: /cgn2_6/ptodata/1/pna/US6016 COMB.seq.*
61: /cgn2_6/ptodata/1/pna/US6017 COMB.seq.*
62: /cgn2_6/ptodata/1/pna/US6018 COMB.seq.*
63: /cgn2_6/ptodata/1/pna/US6019 COMB.seq.*
64: /cgn2_6/ptodata/1/pna/US6020 COMB.seq.*
65: /cgn2_6/ptodata/1/pna/US6021 COMB.seq.*
66: /cgn2_6/ptodata/1/pna/US6022 COMB.seq.*
67: /cgn2_6/ptodata/1/pna/US6023 COMB.seq.*
68: /cgn2_6/ptodata/1/pna/US6024 COMB.seq.*
69: /cgn2_6/ptodata/1/pna/US6025 COMB.seq.*
70: /cgn2_6/ptodata/1/pna/US6026 COMB.seq.*
71: /cgn2_6/ptodata/1/pna/US6027 COMB.seq.*
72: /cgn2_6/ptodata/1/pna/US6028 COMB.seq.*
73: /cgn2_6/ptodata/1/pna/US6029 COMB.seq.*
74: /cgn2_6/ptodata/1/pna/US6030 COMB.seq.*
75: /cgn2_6/ptodata/1/pna/US6031 COMB.seq.*
76: /cgn2_6/ptodata/1/pna/US6032 COMB.seq.*
77: /cgn2_6/ptodata/1/pna/US6033 COMB.seq.*
78: /cgn2_6/ptodata/1/pna/US6034 COMB.seq.*
79: /cgn2_6/ptodata/1/pna/US6035 COMB.seq.*
80: /cgn2_6/ptodata/1/pna/US6036 COMB.seq.*
81: /cgn2_6/ptodata/1/pna/US6037 COMB.seq.*
82: /cgn2_6/ptodata/1/pna/US6038 COMB.seq.*
83: /cgn2_6/ptodata/1/pna/US6039 COMB.seq.*
84: /cgn2_6/ptodata/1/pna/US6040 COMB.seq.*
85: /cgn2_6/ptodata/1/pna/US6041 COMB.seq.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	19.4	88.2	231	21	US-09-540-229-481
2	19.4	88.2	237	13	US-08-903-802-1160
3	19.4	88.2	237	21	US-09-540-210B-29243
4	19.4	88.2	270	13	US-08-993-774-322
5	19.4	88.2	270	20	US-09-539-800-7233
6	19.4	88.2	270	20	US-09-539-800B-7233
7	19.4	88.2	270	20	US-09-539-800C-7233
8	19.4	88.2	270	47	US-60-034-975-322
9	19.4	88.2	321	25	US-09-540-208-54639
10	19.4	88.2	345	21	US-09-540-208-54639
11	19.4	88.2	384	17	US-09-399-932-3999
12	19.4	88.2	408	1	PCT-US01-01307-369
13	19.4	88.2	408	39	US-10-092-302-369
14	19.4	88.2	417	19	US-09-528-409-48934
15	19.4	88.2	417	35	US-09-933-524-48934
16	19.4	88.2	417	35	US-09-933-524A-48934
17	19.4	88.2	448	29	US-09-721-589-51
18	19.4	88.2	458	5	US-08-196-362A-7619
19	19.4	88.2	458	5	US-08-196-362D-7619
20	19.4	88.2	458	5	US-08-196-362E-7619
21	19.4	88.2	458	5	US-08-196-362F-7619
					Sequence 481, App
					Sequence 1160, App
					Sequence 29243, A
					Sequence 322, App
					Sequence 7233, App
					Sequence 7233, App
					Sequence 7233, App
					Sequence 322, App
					Sequence 2566, App
					Sequence 4839, A
					Sequence 369, App
					Sequence 369, App
					Sequence 48934, A
					Sequence 48934, A
					Sequence 51, App1
					Sequence 7619, App
					Sequence 7619, App
					Sequence 7619, App

Mon Apr 21 10:30:26 2003

us-09-270-437d-13.rnppm

```

22 19.4 88.2 458 5 US-08-196-362G-7619
23 19.4 88.2 465 17 US-09-304-649-680
24 19.4 88.2 465 32 US-09-840-424-680
25 19.4 88.2 467 35 US-09-528-409-49810
26 19.4 88.2 467 35 US-09-933-524-49810
27 19.4 88.2 467 35 US-09-933-524A-49810
28 19.4 88.2 546 56 US-60-125-787-430
29 19.4 88.2 598 62 PCT-US01-08631-6785
30 19.4 88.2 2010 1 PCT-US01-08631-6785
31 19.4 88.2 2935 8 US-08-422-420-1
32 19.4 88.2 2935 15 US-09-149-988-1
33 19.4 88.2 2935 34 US-09-927-180-1
34 19.4 88.2 3283 16 US-09-270-437-8
35 19.4 88.2 3283 33 US-09-899-651-8
36 19.4 88.2 3374 1 PCT-US01-01307-113
37 19.4 88.2 3374 39 US-10-092-302-113
38 19.4 88.2 3412 16 US-09-270-437-6
39 19.4 88.2 3412 33 US-09-899-651-6
40 19.4 88.2 3650 61 US-60-172-360-27652
41 19.4 88.2 3667 1 PCT-US01-08631-6786
42 19.4 88.2 3667 39 US-10-097-340-146
43 19.4 88.2 3667 84 US-60-406-385-48
44 19.4 88.2 3694 1 PCT-US01-01341-329
45 19.4 88.2 3694 1 PCT-US01-01341-329

```

ALIGNMENTS

```

RESULT 1
US-09-540-229-481
; Sequence 481, Application US/09540229
; GENERAL INFORMATION:
; APPLICANT: Seilhamer, Jeffrey J.
; APPLICANT: Delegeane, Angelo M.
; APPLICANT: Stuart, Susan G.
; APPLICANT: Mullahy, Sara J.
; APPLICANT: Naughton, Rebecca E.
; TITLE OF INVENTION: POLYNUCLEOTIDES OF NERVOUS SYSTEM AND SENSORY ORGANS
; FILE REFERENCE: PD-1033 CIP
; CURRENT APPLICATION NUMBER: US/09/540,229
; CURRENT FILING DATE: 2000-03-31
; Prior application data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 193582
; SOFTWARE: PERL Program
; SEQ ID NO 481
; LENGTH: 231
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No: hu00195332
; NAME/KEY: unsure
; LOCATION: 2, 11, 23-24, 33, 35, 39, 54, 94-95, 101, 113-114, 173, 182, 194, 198,
; OTHER INFORMATION: a, t, c, g, or other
US-09-540-229-481

```

```

Query Match 88.2%; Score 19.4; DB 21; Length 231;
Best Local Similarity 95.2%; Pred. No. 1.3e+02;
Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GCTCTTTGGGACAGGAAGT 21
Db 119 GCTCTTTGGGACAGGAAGT 139

```

```

RESULT 2
US-08-903-802-1160
; Sequence 1160, Application US/08903802
; GENERAL INFORMATION:
; APPLICANT: Gooding, Douglas H.
; APPLICANT: Stuve, Laura L.

```

```

; APPLICANT: Stuart, Susan G.
; APPLICANT: Ito, Laura Y.
; APPLICANT: Akelblom, Ingrid E.
; APPLICANT: Delegeane, Angelo M.
; APPLICANT: Naughton, Rebecca E.
; APPLICANT: Klingler, Tod M.
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES
; TITLE OF INVENTION: DERIVED FROM HUMAN BLADDER
; NUMBER OF SEQUENCES: 1441
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION NUMBER: US/08/903,802
; FILING DATE: HEREWITH
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/023,308
; FILING DATE: JULY 31, 1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: CERRONE, MICHAEL C.
; REGISTRATION NUMBER: 39,132
; REFERENCE/DOCKET NUMBER: PD-0202 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 855-0555
; TELEFAX: (415) 845-4166
; INFORMATION FOR SEQ ID NO: 1160:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 237 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; IMMEDIATE SOURCE:
; CLONE: 1322001
; US-08-903-802-1160

```

```

Query Match 88.2%; Score 19.4; DB 13; Length 237;
Best Local Similarity 95.2%; Pred. No. 1.3e+02;
Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

QY 1 GCTCTTTGGGACAGGAAGT 21
Db 143 GCTCTTTGGGACAGGAAGT 163

```

```

RESULT 3
US-09-540-210B-29243
; Sequence 29243, Application US/09540210B
; GENERAL INFORMATION:
; APPLICANT: Seilhamer, Jeffrey J.
; APPLICANT: Delegeane, Angelo M.
; APPLICANT: Stuart, Susan G.
; APPLICANT: Stuve, Laura L.
; APPLICANT: Mullahy, Sara J.
; APPLICANT: Naughton, Rebecca E.
; TITLE OF INVENTION: POLYNUCLEOTIDES OF URINARY TRACT TISSUE
; FILE REFERENCE: PD-1037 CIP
; CURRENT APPLICATION NUMBER: US/09/540,210B
; CURRENT FILING DATE: 2002-04-03
; PRIOR APPLICATION NUMBER: 08/972,899
; PRIOR FILING DATE: November 18, 1997
; PRIOR APPLICATION NUMBER: 08/395,244

```

;; PRIOR FILING DATE: February 27, 1995
;; PRIOR APPLICATION NUMBER: 08/722,922
;; PRIOR FILING DATE: September 27, 1996
;; PRIOR APPLICATION NUMBER: 60/005,526
;; PRIOR FILING DATE: September 29, 1995
;; PRIOR APPLICATION NUMBER: 08/824,029
;; PRIOR FILING DATE: March 25, 1997
;; PRIOR APPLICATION NUMBER: 60/014,010
;; PRIOR FILING DATE: March 25, 1996
;; PRIOR APPLICATION NUMBER: 08/826,847
;; PRIOR FILING DATE: April 10, 1997
;; PRIOR APPLICATION NUMBER: 60/015,533
;; PRIOR FILING DATE: April 10, 1996
;; PRIOR APPLICATION NUMBER: 08/903,555
;; PRIOR FILING DATE: July 31, 1997
;; PRIOR APPLICATION NUMBER: 60/023,308
;; PRIOR FILING DATE: July 31, 1996
;; PRIOR APPLICATION NUMBER: 08/862,178
;; PRIOR FILING DATE: May 22, 1997
;; PRIOR APPLICATION NUMBER: 60/018,217
;; PRIOR FILING DATE: May 23, 1996
;; PRIOR APPLICATION NUMBER: 08/881,589
;; PRIOR FILING DATE: June 24, 1997
;; PRIOR APPLICATION NUMBER: 60/021,275
;; PRIOR FILING DATE: June 25, 1996
;; PRIOR APPLICATION NUMBER: 08/903,802
;; PRIOR FILING DATE: July 31, 1997
;; PRIOR APPLICATION NUMBER: 60/023,308
;; PRIOR FILING DATE: July 31, 1996
;; PRIOR APPLICATION NUMBER: 08/905,881
;; PRIOR FILING DATE: August 1, 1997
;; PRIOR APPLICATION NUMBER: 60/025,204
;; PRIOR FILING DATE: August 1, 1996
;; PRIOR APPLICATION NUMBER: 08/903,471
;; PRIOR FILING DATE: July 30, 1997
;; PRIOR APPLICATION NUMBER: 60/025,478
;; PRIOR FILING DATE: July 31, 1996
;; PRIOR APPLICATION NUMBER: 08/903,556
;; PRIOR FILING DATE: July 31, 1997
;; PRIOR APPLICATION NUMBER: 60/025,217
;; PRIOR FILING DATE: August 22, 1996
;; PRIOR APPLICATION NUMBER: 08/937,142
;; PRIOR FILING DATE: September 23, 1997
;; PRIOR APPLICATION NUMBER: 60/026,598
;; PRIOR FILING DATE: September 24, 1996
;; PRIOR APPLICATION NUMBER: 08/960,746
;; PRIOR FILING DATE: October 29, 1997
;; PRIOR APPLICATION NUMBER: 60/030,144
;; PRIOR FILING DATE: October 30, 1996
;; PRIOR APPLICATION NUMBER: 08/826,847
;; PRIOR FILING DATE: April 10, 1997
;; PRIOR APPLICATION NUMBER: 60/015,533
;; PRIOR FILING DATE: April 10, 1996
;; PRIOR APPLICATION NUMBER: 08/755,524
;; PRIOR FILING DATE: November 22, 1996
;; PRIOR APPLICATION NUMBER: 60/007,495
;; PRIOR FILING DATE: November 22, 1995
;; PRIOR APPLICATION NUMBER: 09/021,031
;; PRIOR FILING DATE: February 10, 1998
;; PRIOR APPLICATION NUMBER: 60/039,325
;; PRIOR FILING DATE: February 13, 1997
;; PRIOR APPLICATION NUMBER: 09/035,172
;; PRIOR FILING DATE: March 4, 1998
;; PRIOR APPLICATION NUMBER: 60/040,431
;; PRIOR FILING DATE: March 5, 1997
;; PRIOR APPLICATION NUMBER: 09/041,894
;; PRIOR FILING DATE: March 12, 1998
;; PRIOR APPLICATION NUMBER: 60/040,199
;; PRIOR FILING DATE: March 14, 1997
;; PRIOR APPLICATION NUMBER: 09/050,817
;; PRIOR FILING DATE: March 30, 1998
;; PRIOR APPLICATION NUMBER: 60/043,792
;; PRIOR FILING DATE: April 11, 1997

;; PRIOR APPLICATION NUMBER: 09/074,999
;; PRIOR FILING DATE: May 8, 1998
;; PRIOR APPLICATION NUMBER: 60/048,431
;; PRIOR FILING DATE: May 29, 1997
;; PRIOR APPLICATION NUMBER: 09/107,592
;; PRIOR FILING DATE: June 30, 1998
;; PRIOR APPLICATION NUMBER: 60/052,751
;; PRIOR FILING DATE: July 1, 1997
;; PRIOR APPLICATION NUMBER: 09/094,079
;; PRIOR FILING DATE: June 9, 1998
;; PRIOR APPLICATION NUMBER: 60/049,975
;; PRIOR FILING DATE: June 13, 1997
;; NUMBER OF SEQ ID NOS: 35654
;; SOFTWARE: PERL Program
;; SEQ ID NO 29243
;; LENGTH: 237
;; TYPE: DNA
;; ORGANISM: Homo sapiens
;; FEATURE:
;; NAME/KEY: misc_feature
;; OTHER INFORMATION: Incyte ID No: hu00029997
US-09-540-210B-29243

Query Match 88.2%; Score 19.4; DB 21; Length 237;
Best Local Similarity 95.2%; Pred. No. 1.3e+02;
Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GCTCTTTGGGACAGGAGGT 21
|||||
Db 143 GCTCTTTGGGACAGGAGCT 163
|||||

RESULT 4

US-08-993-774-322
; Sequence 322, Application US/08993774
; GENERAL INFORMATION:
; APPLICANT: Gooding, Douglas H.
; APPLICANT: Stuve, Laura L.
; APPLICANT: Stuart, Susan G.
; APPLICANT: Ito, Laura Y.
; APPLICANT: Akersblom, Ingrid B.
; APPLICANT: Deleageane, Angelo M.
; APPLICANT: Naughton, Rebecca E.
; APPLICANT: Klingler, Tod M.
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM
; TITLE OF INVENTION: MESENTERY TUMOR
; NUMBER OF SEQUENCES: 4086
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/993,774
; FILING DATE:
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: CERRONE, MICHAEL C.
; REGISTRATION NUMBER: 39,132
; REFERENCE/DOCKET NUMBER: PD-0300P
; TELEPHONE: (415) 855-0555
; TELEFAX: (415) 845-4166
; INFORMATION FOR SEQ ID NO: 322:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 270 base pairs

Mon Apr 21 10:30:26 2003

us-09-270-437d-13.rnmpm

TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
IMMEDIATE SOURCE:
CLONE: 2502822H1
US-08-993-774-322

Query Match 88.2%; Score 19.4; DB 13; Length 270;
Best Local Similarity 95.2%; Pred. No. 1.3e+02;
Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 1 GCTCTTTGGGACAGGAAGCT 21

Db 111 GCTCTTTGGGACAGGAAGCT 131

RESULT 5

US-09-539-800-7233

; Sequence 7233, Application US/09539800

; GENERAL INFORMATION:

; APPLICANT: Seilhamer, Jeffrey J.

; APPLICANT: Delegeane, Angelo M.

; APPLICANT: Stuart, Susan G.

; APPLICANT: Stuve, Laura L.

; APPLICANT: Mullahy, Sara J.

; APPLICANT: Naughton, Rebecca E.

; TITLE OF INVENTION: POLYNUCLEOTIDES OF CONNECTIVE TISSUE

; FILE REFERENCE: PD-1023 CIP

; CURRENT APPLICATION NUMBER: US/09/539,800

; PRIOR FILING DATE: 2000-03-30

; PRIOR FILING DATE: August 16, 1995

; PRIOR FILING DATE: August 16, 1995

; PRIOR FILING DATE: June 27, 1994

; PRIOR FILING DATE: November 4, 1994

; PRIOR FILING DATE: November 4, 1994

; PRIOR FILING DATE: October 3, 1997

; PRIOR FILING DATE: October 3, 1997

; PRIOR FILING DATE: October 4, 1996

; PRIOR FILING DATE: October 4, 1996

; PRIOR FILING DATE: October 4, 1997

; PRIOR FILING DATE: October 4, 1996

; PRIOR FILING DATE: October 4, 1996

; PRIOR FILING DATE: December 18, 1997

; PRIOR FILING DATE: December 18, 1997

; PRIOR FILING DATE: December 20, 1996

; PRIOR FILING DATE: December 20, 1996

; PRIOR FILING DATE: February 10, 1999

; PRIOR FILING DATE: February 10, 1999

; PRIOR FILING DATE: February 12, 1998

; PRIOR FILING DATE: February 12, 1998

; PRIOR FILING DATE: December 1, 1999

; PRIOR FILING DATE: December 1, 1999

; PRIOR FILING DATE: December 10, 1998

; NUMBER OF SEQ ID NOS: 19698

; SOFTWARE: PERL Program

; SEQ ID NO 7233

; LENGTH: 270

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: misc feature

; OTHER INFORMATION: Incyte ID No: hu00328727

US-09-539-800-7233

Query Match 88.2%; Score 19.4; DB 20; Length 270;

Best Local Similarity 95.2%; Pred. No. 1.3e+02;

Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 1 GCTCTTTGGGACAGGAAGCT 21

Db 111 GCTCTTTGGGACAGGAAGCT 131

RESULT 6

US-09-539-800B-7233

; Sequence 7233, Application US/09539800B

; GENERAL INFORMATION:

; APPLICANT: Seilhamer, Jeffrey J.

; APPLICANT: Delegeane, Angelo M.

; APPLICANT: Stuart, Susan G.

; APPLICANT: Stuve, Laura L.

; APPLICANT: Mullahy, Sara J.

; APPLICANT: Naughton, Rebecca E.

; TITLE OF INVENTION: POLYNUCLEOTIDES OF CONNECTIVE TISSUE

; FILE REFERENCE: PD-1023 CIP

; CURRENT APPLICATION NUMBER: US/09/539,800B

; PRIOR FILING DATE: 2000-03-30

; PRIOR FILING DATE: August 16, 1995

; PRIOR FILING DATE: August 16, 1995

; PRIOR FILING DATE: June 27, 1994

; PRIOR FILING DATE: November 4, 1994

; PRIOR FILING DATE: November 4, 1994

; PRIOR FILING DATE: October 3, 1997

; PRIOR FILING DATE: October 3, 1997

; PRIOR FILING DATE: October 4, 1996

; PRIOR FILING DATE: October 4, 1996

; PRIOR FILING DATE: October 4, 1997

; PRIOR FILING DATE: October 4, 1996

; PRIOR FILING DATE: October 4, 1996

; PRIOR FILING DATE: December 18, 1997

; PRIOR FILING DATE: December 18, 1997

; PRIOR FILING DATE: December 20, 1996

; PRIOR FILING DATE: December 20, 1996

; PRIOR FILING DATE: February 10, 1999

; PRIOR FILING DATE: February 10, 1999

; PRIOR FILING DATE: February 12, 1998

; PRIOR FILING DATE: February 12, 1998

; PRIOR FILING DATE: December 1, 1999

; PRIOR FILING DATE: December 1, 1999

; PRIOR FILING DATE: December 10, 1998

; NUMBER OF SEQ ID NOS: 19698

; SOFTWARE: PERL Program

; SEQ ID NO 7233

; LENGTH: 270

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: misc feature

; OTHER INFORMATION: Incyte ID No: hu00328727

US-09-539-800B-7233

Query Match 88.2%; Score 19.4; DB 20; Length 270;

Best Local Similarity 95.2%; Pred. No. 1.3e+02;

Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 1 GCTCTTTGGGACAGGAAGCT 21

Db 111 GCTCTTTGGGACAGGAAGCT 131

RESULT 7

US-09-539-800C-7233

; Sequence 7233, Application US/09539800C

; GENERAL INFORMATION:

; APPLICANT: Seilhamer, Jeffrey J.

; APPLICANT: Delegeane, Angelo M.

; APPLICANT: Stuart, Susan G.

; APPLICANT: Stuve, Laura L.

; APPLICANT: Mullahy, Sara J.

APPLICANT: Naughton, Rebecca E.
TITLE OF INVENTION: POLYNUCLEOTIDES OF CONNECTIVE TISSUE
FILE REFERENCE: PD-1023 CIP
CURRENT APPLICATION NUMBER: US/09/539,800C
CURRENT FILING DATE: 2000-03-30
PRIOR APPLICATION NUMBER: 08/521,383
PRIOR FILING DATE: August 16, 1995
PRIOR APPLICATION NUMBER: 08/271,217
PRIOR FILING DATE: June 27, 1994
PRIOR APPLICATION NUMBER: 08/334,881
PRIOR FILING DATE: November 4, 1994
PRIOR APPLICATION NUMBER: 08/943,978
PRIOR FILING DATE: October 3, 1997
PRIOR APPLICATION NUMBER: 60/028,732
PRIOR FILING DATE: October 4, 1996
PRIOR APPLICATION NUMBER: 08/943,979
PRIOR FILING DATE: October 4, 1997
PRIOR APPLICATION NUMBER: 60/027,782
PRIOR FILING DATE: October 4, 1996
PRIOR APPLICATION NUMBER: 08/993,774
PRIOR FILING DATE: December 18, 1997
PRIOR APPLICATION NUMBER: 60/034,975
PRIOR FILING DATE: December 20, 1996
PRIOR APPLICATION NUMBER: 09/250,003
PRIOR FILING DATE: February 10, 1999
PRIOR APPLICATION NUMBER: 60/074,364
PRIOR FILING DATE: February 12, 1998
PRIOR APPLICATION NUMBER: 09/452,747
PRIOR FILING DATE: December 1, 1999
PRIOR APPLICATION NUMBER: 60/111,910
PRIOR FILING DATE: December 10, 1998
NUMBER OF SEQ ID NOS: 19698
SOFTWARE: PERL Program
SEQ ID NO 7233
LENGTH: 270
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
OTHER INFORMATION: Incyte ID No: hu00328727
US-09-539-800C-7233

Query Match 88.2%; Score 19.4; DB 20; Length 270;
Best Local Similarity 95.2%; Pred. No. 1.3e+02;
Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GCTCTTTGGGACAGGAGGT 21
|||||
DB 111 GCTCTTTGGGACAGGAGGT 131

RESULT 8
US-60-034-975-322
Sequence 322, Application US/60034975
GENERAL INFORMATION:
APPLICANT: Gooding, Douglas H.
APPLICANT: Stuve, Laura L.
APPLICANT: Stuart, Susan G.
APPLICANT: Ito, Laura Y.
APPLICANT: Akerblom, Ingrid E.
APPLICANT: Delegeane, Angelo M.
APPLICANT: Naughton, Rebecca E.
APPLICANT: Klingler, Tod M.
TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM
TITLE OF INVENTION: MESENTERY TUMOR
NUMBER OF SEQUENCES: 4086
CORRESPONDENCE ADDRESS:
ADDRESSER: INCYTE PHARMACEUTICALS, INC.
STREET: 3174 PORTER DRIVE
CITY: PALO ALTO
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94304

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/60/034,975
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: CERRONE, MICHAEL C.
REGISTRATION NUMBER: 39,132
REFERENCE/DOCKET NUMBER: PD-0300P
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 855-0555
TELEFAX: (415) 845-4166
INFORMATION FOR SEQ ID NO: 322:
SEQUENCE CHARACTERISTICS:
LENGTH: 270 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
IMMEDIATE SOURCE:
CLONE: 2502822H1
US-60-034-975-322

Query Match 88.2%; Score 19.4; DB 47; Length 270;
Best Local Similarity 95.2%; Pred. No. 1.3e+02;
Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GCTCTTTGGGACAGGAGGT 21
|||||
DB 111 GCTCTTTGGGACAGGAGGT 131

RESULT 9
US-09-644-871-2566
Sequence 2566, Application US/09644871
GENERAL INFORMATION:
APPLICANT: Holtzman, Douglas A.
APPLICANT: Pan, Yang
TITLE OF INVENTION: NOVEL NUCLEIC ACID MOLECULES AND USES
FILE REFERENCE: 1600.1167-001
CURRENT APPLICATION NUMBER: US/09/644,871
CURRENT FILING DATE: 2000-08-28
PRIOR APPLICATION NUMBER: 60/151,059
PRIOR FILING DATE: 1999-08-27
NUMBER OF SEQ ID NOS: 9739
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 2566
LENGTH: 321
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
LOCATION: (1)...(321)
OTHER INFORMATION: n = A,T,C or G
US-09-644-871-2566

Query Match 88.2%; Score 19.4; DB 25; Length 321;
Best Local Similarity 95.2%; Pred. No. 1.3e+02;
Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GCTCTTTGGGACAGGAGGT 21
|||||
DB 122 GCTCTTTGGGACAGGAGGT 142

RESULT 10
US-09-540-208-54639
Sequence 54639, Application US/09540208

Mon Apr 21 10:30:26 2003

us-09-270-437d-13.rnnp

```

; GENERAL INFORMATION:
; APPLICANT: Seilhamer, Jeffrey J.
; APPLICANT: Deleageane, Angelo M.
; APPLICANT: Stuart, Susan G.
; APPLICANT: Stuve, Laura L.
; APPLICANT: Mullahy, Sara J.
; APPLICANT: Naughton, Rebecca E.
; TITLE OF INVENTION: POLYNUCLEOTIDES OF MALE REPRODUCTIVE TISSUE
; FILE REFERENCE: PD-1029 CIP
; CURRENT APPLICATION NUMBER: US/09/540,208
; CURRENT FILING DATE: 2000-03-31
; Prior application data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 70811
; SOFTWARE: PERL Program
; SEQ ID NO 54639
; LENGTH: 345
; TYPE: DNA
; ORGANISM: Homo sapiens
; NAME/KEY: misc feature
; OTHER INFORMATION: incyte ID No: hu01153334
; US-09-540-208-54639

Query Match      88.2%; Score 19.4; DB 21; Length 345;
Best Local Similarity 95.2%; Pred. No. 1.4e+02;
Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GCTCTTTGGGACAGGAGGT 21
    |||||
Db 151 GCTCTTTGGGACAGGAGCT 171

RESULT 11
US-09-399-932-3999
; Sequence 3999, Application US/09399932
; GENERAL INFORMATION:
; APPLICANT: Holtzman, Douglas A.
; APPLICANT: Gearing, David P.
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES DERIVED FROM A
; TITLE OF INVENTION: HUMAN BURKITT'S LYMPHOMA LIBRARY
; FILE REFERENCE: MLN98-47PM
; CURRENT APPLICATION NUMBER: US/09/399,932
; CURRENT FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: 60/100,464
; PRIOR FILING DATE: 1998-09-15
; PRIOR APPLICATION NUMBER: 60/101,670
; PRIOR FILING DATE: 1998-09-24
; PRIOR APPLICATION NUMBER: 60/106,456
; PRIOR FILING DATE: 1998-10-30
; PRIOR APPLICATION NUMBER: 60/126,902
; PRIOR FILING DATE: 1999-03-30
; PRIOR APPLICATION NUMBER: 60/132,067
; PRIOR FILING DATE: 1999-04-30
; PRIOR APPLICATION NUMBER: 60/144,447
; PRIOR FILING DATE: 1999-07-16
; NUMBER OF SEQ ID NOS: 5743
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 3999
; LENGTH: 384
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-399-932-3999

Query Match      88.2%; Score 19.4; DB 17; Length 384;
Best Local Similarity 95.2%; Pred. No. 1.4e+02;
Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GCTCTTTGGGACAGGAGGT 21
    |||||
Db 130 GCTCTTTGGGACAGGAGCT 150

RESULT 12
; Sequence 369, Application PC/TUS0101307
; GENERAL INFORMATION:
; APPLICANT: Human Genome Sciences, Inc., et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PTZ18PCT
; CURRENT APPLICATION NUMBER: PCT/US01/01307
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1040
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 369
; LENGTH: 408
; TYPE: DNA
; ORGANISM: Homo sapiens
; NAME/KEY: SITE
; LOCATION: (124)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (328)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (337)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (363)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (385)
; OTHER INFORMATION: n equals a,t,g, or c
; PCT-US01-01307-369

Query Match      88.2%; Score 19.4; DB 1; Length 408;
Best Local Similarity 95.2%; Pred. No. 1.4e+02;
Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GCTCTTTGGGACAGGAGGT 21
    |||||
Db 65 GCTCTTTGGGACAGGAGCT 85

RESULT 13
US-10-092-302-369
; Sequence 369, Application US/10092302
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PTZ18C1
; CURRENT APPLICATION NUMBER: US/10/092,302
; CURRENT FILING DATE: 2002-03-07
; NUMBER OF SEQ ID NOS: 1040
; Prior application removed - See File Wrapper or Palm
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 369
; LENGTH: 408
; TYPE: DNA
; ORGANISM: Homo sapiens
; NAME/KEY: misc feature
; LOCATION: (124)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (328)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (337)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (363)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (385)
```



```

; OTHER INFORMATION: n equals a,t,g, or c
US-10-092-302-369

Query Match      88.2%; Score 19.4; DB 39; Length 408;
Best Local Similarity 95.2%; Pred.No. 1.4e+02;
Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GCTCTTTGGGACAGGAAGCT 21
    |||||
Db 65 GCTCTTTGGGACAGGAAGCT 85
    |||||

```

```

RESULT 14
US-09-528-409-48934
; Sequence 48934, Application US/09528409
; GENERAL INFORMATION:
; APPLICANT: Drmanac, Radoje T.
; APPLICANT: Labat, Ivan
; APPLICANT: Stache-Crain, Birgit
; APPLICANT: Dickson, Mark
; APPLICANT: Jones, Lee W.
; TITLE OF INVENTION: Novel Nucleic Acid Sequences Obtained
; TITLE OF INVENTION: From Various Libraries
; FILE REFERENCE: 774
; CURRENT APPLICATION NUMBER: US/09/528,409
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: 60/125,453
; NUMBER OF SEQ ID NOS: 116231
; SOFTWARE: Hy-patent.pl Version 3.1
; SEQ ID NO 48934
; LENGTH: 417
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(417)
; OTHER INFORMATION: n = A,T,C or G
US-09-528-409-48934

```

```

Query Match      88.2%; Score 19.4; DB 19; Length 417;
Best Local Similarity 95.2%; Pred.No. 1.4e+02;
Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GCTCTTTGGGACAGGAAGCT 21
    |||||
Db 117 GCTCTTTGGGACAGGAAGCT 137
    |||||

```

```

RESULT 15
US-09-933-524-48934
; Sequence 48934, Application US/09933524
; GENERAL INFORMATION:
; APPLICANT: Drmanac, Radoje T.
; APPLICANT: Labat, Ivan
; APPLICANT: Stache-Crain, Birgit
; APPLICANT: Dickson, Mark
; APPLICANT: Jones, Lee W.
; TITLE OF INVENTION: Novel Nucleic Acid Sequences Obtained
; TITLE OF INVENTION: From Various Libraries
; FILE REFERENCE: 774
; CURRENT APPLICATION NUMBER: US/09/933,524
; PRIOR FILING DATE: 2001-08-20
; PRIOR APPLICATION NUMBER: 09/528,409
; PRIOR FILING DATE: 2000-03-17
; NUMBER OF SEQ ID NOS: 116231
; SOFTWARE: Hy-patent.pl Version 3.1
; SEQ ID NO 48934
; LENGTH: 417
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature

```

```

; LOCATION: (1)...(417)
; OTHER INFORMATION: n = A,T,C or G
US-09-933-524-48934

Query Match      88.2%; Score 19.4; DB 35; Length 417;
Best Local Similarity 95.2%; Pred.No. 1.4e+02;
Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

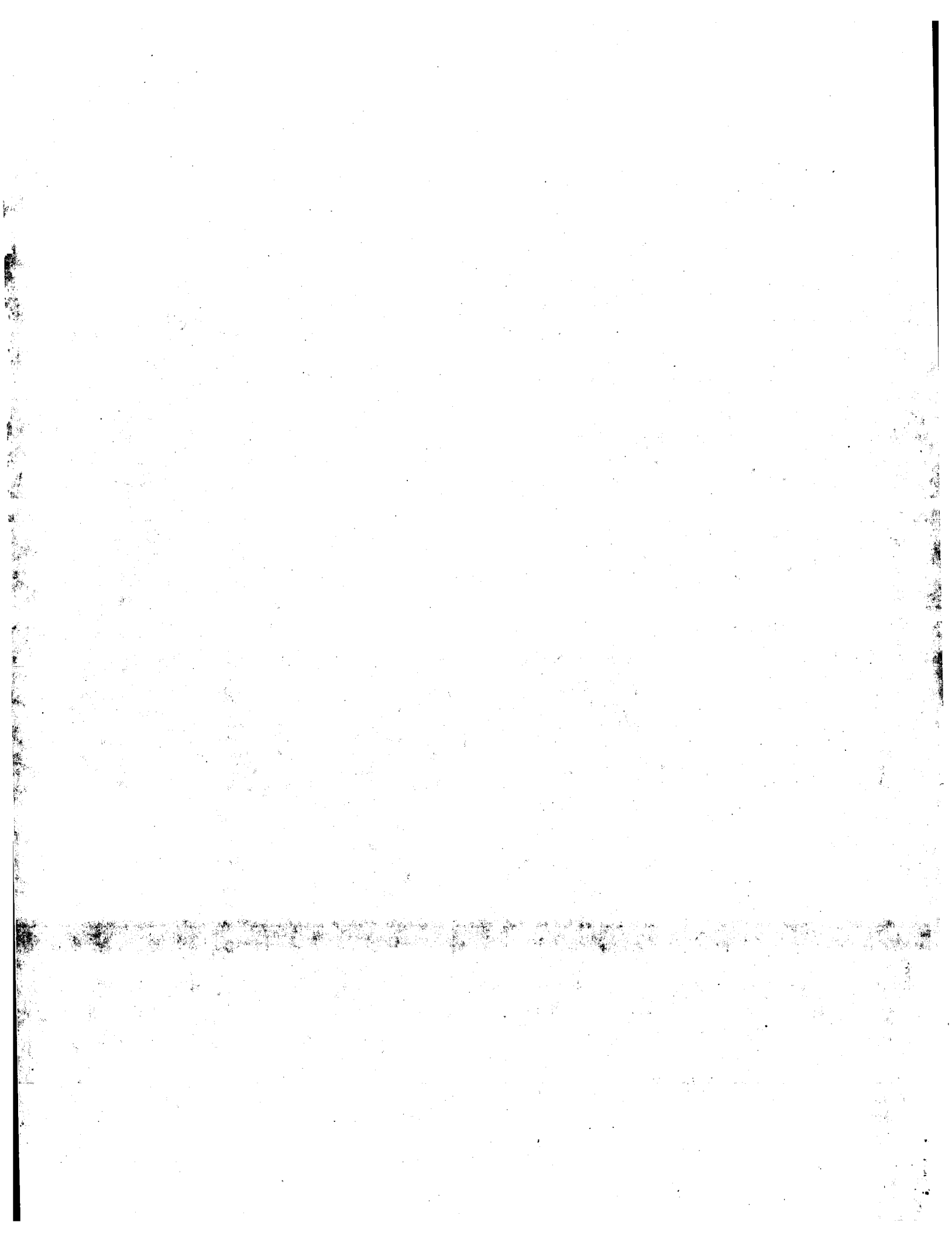
Qy 1 GCTCTTTGGGACAGGAAGCT 21
    |||||
Db 117 GCTCTTTGGGACAGGAAGCT 137
    |||||

```

```

Search completed: April 18, 2003, 09:35:47
Job time : 1212 secs

```



GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 18, 2003, 06:35:28 ; Search time 245.167 Seconds
(without alignment)
400.770 Million cell updates/sec

Title: US-09-270-437D-13

Perfect score: 22
Sequence: 1 gctctttgggacaggaagtc 22

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 5897297 seqs, 2233080881 residues

Total number of hits satisfying chosen parameters: 11794594

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Pending Patents NA New:
1: /cgn2_6/ptodata/1/pna/US06_NEW_COMB.seq.*
2: /cgn2_6/ptodata/1/pna/US06_NEW_COMB.seq.*
3: /cgn2_6/ptodata/1/pna/US07_NEW_COMB.seq.*
4: /cgn2_6/ptodata/1/pna/US08_NEW_COMB.seq.*
5: /cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq.*
6: /cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq.*
7: /cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq.*
8: /cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq.*
9: /cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq.*
10: /cgn2_6/ptodata/1/pna/US60_NEW_COMB.seq.*
11: /cgn2_6/ptodata/1/pna/US60_NEW_COMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	22	100.0	22	5	US-09-270-437D-13
2	19.4	88.2	364	6	US-09-513-999C-3265
3	19.4	88.2	458	6	US-09-912-293-228701
4	19.4	88.2	871	6	US-09-724-676-8826
5	19.4	88.2	871	6	US-09-724-676A-8826
6	19.4	88.2	1707	1	PCT-US02-31357-39
7	19.4	88.2	1707	9	US-10-262-445-39
8	19.4	88.2	3237	8	US-10-170-235-5179
9	19.4	88.2	3237	11	US-60-452-680-612
10	19.4	88.2	3237	11	US-60-453-135-482
11	19.4	88.2	3237	11	US-60-453-050-482
12	19.4	88.2	3283	5	US-09-270-437D-8
13	19.4	88.2	3412	5	US-09-270-437D-6
14	19.4	88.2	3427	8	US-10-170-235-21870
15	19.4	88.2	3427	11	US-60-453-135-3834
16	19.4	88.2	3427	11	US-60-453-050-3834
17	19.4	88.2	3445	11	US-60-452-680-3237
18	19.4	88.2	3475	8	US-10-170-235-22086
19	19.4	88.2	3475	11	US-60-453-135-3835
20	19.4	88.2	3475	11	US-60-453-050-3835
21	18.8	85.5	30922	7	US-09-949-016-16700
22	18.4	83.6	598359	7	US-09-947-911-215

23	18	81.8	84839	7	US-09-949-016-15816	Sequence 15816, A
c 24	17.8	80.9	9968	7	US-09-949-016-12472	Sequence 12472, A
c 25	17.8	80.9	9969	7	US-09-949-016-15903	Sequence 15903, A
c 26	17.4	79.1	403	6	US-09-912-293-96328	Sequence 96328, A
c 27	17.4	79.1	474	6	US-09-912-293-3245	Sequence 3245, Ap
c 28	17.4	79.1	4101	9	US-10-144-771-20672	Sequence 20672, A
c 29	17.2	78.2	201	10	US-60-455-444-42575	Sequence 42575, A
c 30	17.2	78.2	201	11	US-60-453-135-70940	Sequence 70940, A
c 31	17.2	78.2	201	11	US-60-453-050-70940	Sequence 70940, A
c 32	17.2	78.2	442	6	US-09-534-850-18297	Sequence 18297, A
c 33	17.2	78.2	2580	9	US-10-092-900A-153	Sequence 153, App
c 34	17.2	78.2	2890	8	US-10-170-235-32027	Sequence 32027, A
c 35	17.2	78.2	2890	11	US-60-441-839-137	Sequence 137, App
c 36	17.2	78.2	2890	11	US-60-452-680-9301	Sequence 9301, Ap
c 37	17.2	78.2	2890	11	US-60-453-135-5460	Sequence 5460, Ap
c 38	17.2	78.2	2890	11	US-60-453-050-5460	Sequence 5460, Ap
c 39	17.2	78.2	3560	9	US-10-338-044-2260	Sequence 2260, Ap
c 40	17.2	78.2	4933	1	PCT-US03-05336-52	Sequence 52, Appl
c 41	17.2	78.2	4946	8	US-10-170-235-40560	Sequence 40560, A
c 42	17.2	78.2	4985	7	US-09-949-016-152	Sequence 152, App
c 43	17.2	78.2	6532	8	US-10-170-235-23701	Sequence 23701, A
c 44	17.2	78.2	6532	10	US-60-455-444-3409	Sequence 3409, Ap
c 45	17.2	78.2	6532	11	US-60-453-135-6603	Sequence 6603, Ap

ALIGNMENTS

RESULT 1
US-09-270-437D-13
; Sequence 13, Application US/09270437D
; GENERAL INFORMATION:
; APPLICANT: Chen, Yao-Tsang
; APPLICANT: Gure, Ali
; APPLICANT: Tsang, Solam
; APPLICANT: Stockert, Elisabeth
; APPLICANT: Jager, Elke
; APPLICANT: Knuth, Alexander
; APPLICANT: Old, Lloyd J.
; TITLE OF INVENTION: Isolated Nucleic Acid Molecules Encoding Cancer Associated Antigen
; FILE REFERENCE: LUD 5538.1
; CURRENT APPLICATION NUMBER: US/09/270,437D
; PRIOR FILING DATE: 1999-03-16
; PRIOR APPLICATION NUMBER: 09/061,709
; NUMBER OF SEQ ID NOS: 23
; SEQ ID NO 13
; LENGTH: 22
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
US-09-270-437D-13

Query Match 100.0%; Score 22; DB 5; Length 22;
Best Local Similarity 100.0%; Pred. No. 1.4;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 GCTCTTTGGGACAGGAAGTC 22
Db 1 GCTCTTTGGGACAGGAAGTC 22

RESULT 2
US-09-513-999C-3265
; Sequence 3265, Application US/09513999C
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Duclert, A.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; FILE REFERENCE: 59.US02.REG
; CURRENT APPLICATION NUMBER: US/09/513,999C

PRIOR FILING DATE: 2001-03-21

FEATURE:
NAME/KEY: misc feature
LOCATION: (403)..(403)
OTHER INFORMATION: n is equal to a,t,g, or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (410)..(410)
OTHER INFORMATION: n is equal to a,t,g, or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (416)..(416)
OTHER INFORMATION: n is equal to a,t,g, or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (419)..(419)
OTHER INFORMATION: n is equal to a,t,g, or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (428)..(428)
OTHER INFORMATION: n is equal to a,t,g, or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (435)..(435)
OTHER INFORMATION: n is equal to a,t,g, or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (457)..(457)
OTHER INFORMATION: n is equal to a,t,g, or c
US-09-912-293-228701

Query Match 88.2%; Score 19.4; DB 6; Length 458;
Best Local Similarity 95.2%; Pred. No. 31;
Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GCTCTTTGGGACAGGAAGGT 21
| | | | | | | | | | | | | | | | | | | | | |
Db 150 GCTCTTTGGGACAGGAAGCT 170

RESULT 4
US-09-724-676-8826/c
Sequence 8826, Application US/09724676
GENERAL INFORMATION:
APPLICANT: Compugen LTD
TITLE OF INVENTION: Variants of alternative splicing
FILE REFERENCE: 129181.4 Compugen
CURRENT APPLICATION NUMBER: US/09/724,676
CURRENT FILING DATE: 2000-11-28
NUMBER OF SEQ ID NOS: 97222
SOFTWARE: PatentIn version 3.2
SEQ ID NO 8826
LENGTH: 871
TYPE: DNA
ORGANISM: Homo sapiens
US-09-724-676-8826

Query Match 88.2%; Score 19.4; DB 6; Length 871;
Best Local Similarity 95.2%; Pred. No. 33;
Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GCTCTTTGGGACAGGAAGGT 21
| | | | | | | | | | | | | | | | | | | | | |
Db 374 GCTCTTTGGGACAGGAAGCT 354

RESULT 5
US-09-724-676A-8826/c
Sequence 8826, Application US/09724676A
GENERAL INFORMATION:
APPLICANT: Compugen LTD
TITLE OF INVENTION: Variants of alternative splicing
FILE REFERENCE: 129181.4 Compugen
CURRENT APPLICATION NUMBER: US/09/724,676A

CURRENT FILING DATE: 2000-11-28
NUMBER OF SEQ ID NOS: 97222
SOFTWARE: PatentIn version 3.2
SEQ ID NO 8826
LENGTH: 871
TYPE: DNA
ORGANISM: Homo sapiens
US-09-724-676A-8826

Query Match 88.2%; Score 19.4; DB 6; Length 871;
Best Local Similarity 95.2%; Pred. No. 33;
Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GCTCTTTGGGACAGGAAGGT 21
| | | | | | | | | | | | | | | | | | | | | |
Db 374 GCTCTTTGGGACAGGAAGCT 354

RESULT 6
PCT-US02-31357-39
Sequence 39, Application PC/TUS0231357
GENERAL INFORMATION:
APPLICANT: Curagen Corporation, et al
TITLE OF INVENTION: NOVEL HUMAN PROTEINS, POLYNUCLEOTIDES ENCODING THEM AND METHODS OF
FILE REFERENCE: 21402-462D-061
CURRENT APPLICATION NUMBER: PCT/US02/31357
CURRENT FILING DATE: 2002-10-02
PRIOR APPLICATION NUMBER: 60/327,454
PRIOR FILING DATE: 2001-10-05
PRIOR APPLICATION NUMBER: 60/327,917
PRIOR FILING DATE: 2001-10-09
PRIOR APPLICATION NUMBER: 60/328,029
PRIOR FILING DATE: 2001-10-09
PRIOR APPLICATION NUMBER: 60/328,056
PRIOR FILING DATE: 2001-10-09
PRIOR APPLICATION NUMBER: 60/328,849
PRIOR FILING DATE: 2001-10-12
PRIOR APPLICATION NUMBER: 60/329,414
PRIOR FILING DATE: 2001-10-15
PRIOR APPLICATION NUMBER: 60/330,142
PRIOR FILING DATE: 2001-10-17
PRIOR APPLICATION NUMBER: 60/341,058
PRIOR FILING DATE: 2001-10-22
PRIOR APPLICATION NUMBER: 60/343,629
PRIOR FILING DATE: 2001-10-24
PRIOR APPLICATION NUMBER: 60/349,575
PRIOR FILING DATE: 2001-10-29
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 133
SOFTWARE: CuraSeqList version 0.1
SEQ ID NO 39
LENGTH: 1707
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (5)..(1669)
PCT-US02-31357-39

Query Match 88.2%; Score 19.4; DB 1; Length 1707;
Best Local Similarity 95.2%; Pred. No. 36;
Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GCTCTTTGGGACAGGAAGGT 21
| | | | | | | | | | | | | | | | | | | | | |
Db 67 GCTCTTTGGGACAGGAAGCT 87

RESULT 7
US-10-262-445-39
Sequence 39, Application US/10262445
GENERAL INFORMATION:

```

; TITLE OF INVENTION: TRANSCRIPTS, FOR DETECTING EXPRESSION AND OTHER USES THEREOF
; FILE REFERENCE: CL001380
; CURRENT APPLICATION NUMBER: US/10/170,235
; CURRENT FILING DATE: 2003-03-17
; NUMBER OF SEQ ID NOS: 42514
; SEQ ID NO 5179
; LENGTH: 3237
; TYPE: DNA
; ORGANISM: HUMAN
US-10-170-235-5179

Query Match 88.2%; Score 19.4; DB 8; Length 3237;
Best Local Similarity 95.2%; Pred. No. 38;
Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GCTCTTTGGGGACAGGAAGGT 21
DB 76 GCTCTTTGGGGACAGGAAGCT 96

RESULT 9
US-60-452-680-612
; Sequence 612, Application US/60452680
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; APPLICANT: GRUPE, Andrew
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; FILE REFERENCE: ALZHEIMER'S DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001450
; CURRENT APPLICATION NUMBER: US/60/452,680
; CURRENT FILING DATE: 2003-03-07
; NUMBER OF SEQ ID NOS: 116213
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 612
; LENGTH: 3237
; TYPE: DNA
; ORGANISM: Homo sapiens
US-60-452-680-612

Query Match 88.2%; Score 19.4; DB 11; Length 3237;
Best Local Similarity 95.2%; Pred. No. 38;
Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GCTCTTTGGGGACAGGAAGGT 21
DB 76 GCTCTTTGGGGACAGGAAGCT 96

RESULT 10
US-60-453-135-482
; Sequence 482, Application US/60453135
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; APPLICANT: IAKOUBOVA, Olga
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; FILE REFERENCE: MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001456
; CURRENT APPLICATION NUMBER: US/60/453,135
; CURRENT FILING DATE: 2003-03-10
; NUMBER OF SEQ ID NOS: 82762
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 482
; LENGTH: 3237
; TYPE: DNA
; ORGANISM: Homo sapiens
US-60-453-135-482

Query Match 88.2%; Score 19.4; DB 11; Length 3237;
Best Local Similarity 95.2%; Pred. No. 38;
Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

Db 76 GCTCTTTGGGACAGGAAGCT 96

RESULT 11

US-60-453-050-482
; Sequence 482, Application US/60453050
; GENERAL INFORMATION:

; APPLICANT: CARGILL, Michele

; APPLICANT: LUKE, May

; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
STENOSIS, METHODS OF DETECTION AND USES THEREOF

; FILE REFERENCE: CL001457

; CURRENT APPLICATION NUMBER: US/60/453,050

; CURRENT FILING DATE: 2003-03-10

; NUMBER OF SEQ ID NOS: 82762

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 482

; LENGTH: 3237

; TYPE: DNA

; ORGANISM: Homo sapiens

US-60-453-050-482

Query Match 88.2%; Score 19.4; DB 11; Length 3237;

Best Local Similarity 95.2%; Pred. No. 38;

Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GCTCTTTGGGACAGGAAGCT 21

Db 76 GCTCTTTGGGACAGGAAGCT 96

RESULT 12

US-09-270-437D-8

; Sequence 8, Application US/09270437D

; GENERAL INFORMATION:

; APPLICANT: Chen, Yao-Tseng

; APPLICANT: Gure, Ali

; APPLICANT: Tsang, Solam

; APPLICANT: Stockert, Elisabeth

; APPLICANT: Jager, Elke

; APPLICANT: Knuth, Alexander

; APPLICANT: Old, Lloyd J.

; TITLE OF INVENTION: Isolated Nucleic Acid Molecules Encoding Cancer Associated Antigen

; FILE REFERENCE: LUD 5538.1

; CURRENT APPLICATION NUMBER: US/09/270,437D

; CURRENT FILING DATE: 1999-03-16

; PRIOR FILING DATE: 09/061,709

; NUMBER OF SEQ ID NOS: 23

; SEQ ID NO 8

; LENGTH: 3283

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: CDS

; LOCATION: 3243

; OTHER INFORMATION: unsure of nucleotide

US-09-270-437D-8

Query Match 88.2%; Score 19.4; DB 5; Length 3283;

Best Local Similarity 95.2%; Pred. No. 38;

Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GCTCTTTGGGACAGGAAGCT 21

Db 132 GCTCTTTGGGACAGGAAGCT 152

RESULT 13

US-09-270-437D-6

; Sequence 6, Application US/09270437D

; GENERAL INFORMATION:

; APPLICANT: Chen, Yao-Tseng
; APPLICANT: Gure, Ali
; APPLICANT: Tsang, Solam
; APPLICANT: Stockert, Elisabeth
; APPLICANT: Jager, Elke
; APPLICANT: Knuth, Alexander
; APPLICANT: Old, Lloyd J.
; TITLE OF INVENTION: Isolated Nucleic Acid Molecules Encoding Cancer Associated Antigen
; FILE REFERENCE: LUD 5538.1
; CURRENT APPLICATION NUMBER: US/09/270,437D
; CURRENT FILING DATE: 1999-03-16
; PRIOR FILING DATE: 09/061,709
; NUMBER OF SEQ ID NOS: 23
; SEQ ID NO 6
; LENGTH: 3412
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:

; NAME/KEY: CDS

; LOCATION: 3372

; OTHER INFORMATION: unsure of nucleotide

US-09-270-437D-6

Query Match 88.2%; Score 19.4; DB 5; Length 3412;

Best Local Similarity 95.2%; Pred. No. 39;

Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GCTCTTTGGGACAGGAAGCT 21

Db 132 GCTCTTTGGGACAGGAAGCT 152

RESULT 14

US-10-170-235-21870

; Sequence 21870, Application US/10170235

; GENERAL INFORMATION:

; APPLICANT: VENTER, J. Craig

; TITLE OF INVENTION: KITS, SUCH AS NUCLEIC ACID ARRAYS, COMPRISING A MAJORITY OF HUMAN
TRANSCRIPTS, FOR DETECTING EXPRESSION AND OTHER USES THEREOF

; FILE REFERENCE: CL001380

; CURRENT APPLICATION NUMBER: US/10/170,235

; CURRENT FILING DATE: 2003-03-17

; NUMBER OF SEQ ID NOS: 42514

; SEQ ID NO 21870

; LENGTH: 3427

; TYPE: DNA

; ORGANISM: HUMAN

US-10-170-235-21870

Query Match 88.2%; Score 19.4; DB 8; Length 3427;

Best Local Similarity 95.2%; Pred. No. 39;

Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GCTCTTTGGGACAGGAAGCT 21

Db 130 GCTCTTTGGGACAGGAAGCT 150

RESULT 15

US-60-453-135-3834

; Sequence 3834, Application US/60453135

; GENERAL INFORMATION:

; APPLICANT: CARGILL, Michele

; APPLICANT: IAKOUBOVA, Olga

; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH

MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF

; FILE REFERENCE: CL001456

; CURRENT APPLICATION NUMBER: US/60/453,135

; CURRENT FILING DATE: 2003-03-10

; NUMBER OF SEQ ID NOS: 82762

; SOFTWARE: FastSeq for Windows Version 4.0

Mon Apr 21 10:30:28 2003

```

; SEQ ID NO 3834
; LENGTH: 3427
; TYPE: DNA
; ORGANISM: Homo sapiens
US-60-453-135-3834

```

```
Query Match      88.2%; Score 19.4; DB 11; Length 3427;
Best Local Similarity 95.2%; Pred. No. 39;
Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

QY . 1 GCTCTTTGGGACAGGAAGT 21
| | | | | | | | | |
Db 130 GCTCTTTGGGACAGGAAGT 150

Search completed: April 18, 2003, 10:50:31
Job time : 249.167 secs

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 18, 2003, 04:47:40 ; Search time 766 Seconds
(without alignment)
465.145 Million cell updates/sec

Title: US-09-270-437D-13

Perfect score: 22
Sequence: 1 gctcttggggacaggaaggtc 22

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

1: em_estba:*

2: em_esthum:*

3: em_estin:*

4: em_estmu:*

5: em_estov:*

6: em_estpl:*

7: em_estro:*

8: em_hcc:*

9: gb_estl:*

10: gb_est2:*

11: gb_hcc:*

12: gb_est3:*

13: gb_est4:*

14: gb_est5:*

15: em_estfun:*

16: em_estom:*

17: gb_gss:*

18: em_gss_hum:*

19: em_gss_inv:*

20: em_gss_pln:*

21: em_gss_vrt:*

22: em_gss_fun:*

23: em_gss_mam:*

24: em_gss_mus:*

25: em_gss_other:*

26: em_gss_pro:*

27: em_gss_rod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	19.4	88.2	330	9	AA226813
2	19.4	88.2	453	9	AL121456
3	19.4	88.2	478	12	BG007027
4	19.4	88.2	479	14	H66979
5	19.4	88.2	620	12	BF914566
6	19.4	88.2	851	14	BQ221568

7	19.4	88.2	874	12	BG481336
8	19.4	88.2	889	12	BG748346
9	18.8	85.5	264	13	BI971125
10	18.8	85.5	309	13	BM484358
11	18.8	85.5	373	12	BF189397
12	18.8	85.5	471	9	AI623653
13	18.4	83.6	566	17	BH261805
14	18.4	83.6	680	17	BH098064
15	18.4	83.6	856	12	BG480515
16	18.4	83.6	2768	11	BC015720
17	18	81.8	738	17	P0222C03U
18	18	81.8	861	17	BQ706255
19	18	81.8	1742	11	AK015621
20	17.8	80.9	299	9	AL597676
21	17.8	80.9	368	10	BH872349
22	17.8	80.9	429	14	Z84036
23	17.8	80.9	434	10	BH863724
24	17.8	80.9	448	10	BH852457
25	17.8	80.9	449	13	BM253624
26	17.8	80.9	458	10	BH852318
27	17.8	80.9	469	10	BH853640
28	17.8	80.9	489	10	BH853310
29	17.8	80.9	528	13	BM107521
30	17.8	80.9	541	14	BQ770242
31	17.8	80.9	582	14	BQ235601
32	17.8	80.9	617	10	BH660288
33	17.8	80.9	649	17	AZ650694
34	17.8	80.9	668	10	BB618315
35	17.8	80.9	673	12	BG768944
36	17.8	80.9	715	17	AG142970
37	17.8	80.9	815	13	BI688088
38	17.8	80.9	1146	13	BM464345
39	17.4	79.1	140	10	BE221716
40	17.4	79.1	150	10	BB163693
41	17.4	79.1	152	10	BB191260
42	17.4	79.1	154	10	BB363001
43	17.4	79.1	154	10	BB438580
44	17.4	79.1	154	10	BB527579
45	17.4	79.1	155	10	BB515762

ALIGNMENTS

RESULT 1
AA226813
LOCUS
DEFINITION
Stratagene NT2 neuronal precursor 937230 Homo sapiens
CDNA clone IMAGE:663785 5', mRNA sequence.
ACCESSION
AA226813
VERSION
AA226813.1
KEYWORDS
EST.
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 330)
Hillier L., Lennon G., Becker M., Bonaldo M.P., Chiapelli B., Chissoe S., Dietrich N., Dubuque T., Favello A., Gish W., Hawkins M., Hultman M., Kucaba T., Lacy M., Le M., Le N., Mardis E., Moore B., Morris M., Parsons J., Prange C., Rifkin L., Rohlfing T., Schellenberg K., Soares M.B., Tan F., Thierry-Mieg J., Travaskis E., Underwood K., Wohlmann P., Waterston R., Wilson R. and Marra M. Generation and analysis of 280,000 human expressed sequence tags
Genome Res. 6 (9), 807-828 (1996)
97044478
Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available royalty-free through LNL ; contact the

Mon Apr 21 10:30:29 2003

IMAGE Consortium (info@image.llnl.gov) for further information.
Seq primer: -28mi3 rev1 ET from Amersham
High quality sequence stop: 321.

FEATURES

source

1. 330
/organism="Homo sapiens"
/db_xref="GDB:5589715"
/db_xref="taxon:9606"
/clone="IMAGE:663785"
/clone_lib="Stratagene NT2 neuronal precursor 937230"
/tissue_type="neuroepithelial cells"
/dev_stage="Ntera-2 neuroepithelial cells"
/lab_host="SOLR (kanamycin resistant)"
/note="Organ: brain; Vector: pBluescript SK-; Site: 1:
ECORI; Site: 2: XhoI; Cloned unidirectionally. Primer:
Oligo dt. Uninduced, exponentially growing neuroepithelial
cells (Ntera-2/ci.D1). Average insert size: 1.0 kb;
Uni-ZAP XR Vector; -5' adaptor sequence: 5' GAATTCGGCAGG
3' -3' adaptor sequence: 5' CTCGAGTTTTTTTTTTTTTTT 3"

BASE COUNT
ORIGIN

Query Match 88.2%; Score 19.4; DB 9; Length 330;
Best Local Similarity 95.2%; Pred. No. 1.8e+02;
Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GCTCTTTGGGACAGGAGGT 21

Db 109 GCTCTTTGGGACAGGAGGT 129

RESULT 2

LOCUS

AL121466 453 bp mRNA linear EST 25-FEB-2000
DEFINITION DKFZp762L097_r1 762 (synonym: hmcl2) Homo sapiens cDNA clone
DKFZp762L097 5', mRNA sequence.

ACCESSION

AL121466.1 GI:5927467

VERSION

EST.

KEYWORDS

SOURCE

human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 453)
Ottewaelder, B., Obermaier, B., Mewes, H.W., Gassenhuber, J. and
Wiemann, S.

TITLE

JOURNAL

COMMENT

Unpublished (1999)
Contact: Ottewaelder B
MIPS
Am Klopferspitz 18a D-82152 Martinsried, Germany
This is the 5' sequence of the clone insert
Clone from S. Wilmann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wilmann@dkfz-heidelberg.de;
Research Center (DKFZ); Email s.wilmann@dkfz-heidelberg.de;
sequencing consortium of the German Genome Project. No s1 sequence
available.
This clone (DKFZp762L097) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.

FEATURES
source

1. 453
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="DKFZp762L097"
/clone_lib="hmcl2" (synonym: hmcl2)
/tissue_type="melanoma (MeWo cell line)"
/dev_stage="adult"
/lab_host="DH10B"
/note="vector: pSport1; Site_1: NotI; Site_2: SalI"

BASE COUNT
ORIGIN

117 a 121 c 137 g 78 t

Query Match

Best Local Similarity

Matches

Conservative

Mismatches

Indels

Gaps

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

LOCUS H66979 479 bp mRNA linear EST 18-OCT-1995
 DEFINITION Yul7a11.r1 Soares fetal liver spleen INFILS Homo sapiens cDNA clone
 IMAGE:234044 5' similar to contains Alu repetitive element; , mRNA
 sequence.
 ACCESSION H66979
 VERSION H66979.1 GI:1025719
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 479)
 AUTHORS Hillier, L., Clark, N., Dubuque, T., Elliston, K., Hawkins, M., Holman
 M., Hultman, M., Kucaba, T., Le, M., Lennon, G., Marra, M., Parsons, J.,
 Rifkin, L., Rohlfing, T., Soares, M., Tan, F., Trevaskis, E., Waterston
 R., Williamson, A., Wohlmann, P. and Wilson, R.
 TITLE The WashU-Merck EST Project
 JOURNAL Unpublished (1995)
 COMMENT Contact: wilson RK
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: est@watson.wustl.edu
 Insert Size: 1511
 High quality sequence stops: 401
 Source: IMAGE Consortium, LLNL
 This clone is available royalty-free through LLNL; contact the
 IMAGE Consortium (info@image.llnl.gov) for further information.
 Insert Length: 1511 Std Error: 0.00
 Seq primer: M13RP1
 High quality sequence stop: 401.
 Location/Qualifiers
 1..479
 /organism="Homo sapiens"
 /db_xref="GDB:378721"
 /db_xref="taxon:9606"
 /clone="IMAGE:234044"
 /clone_lib="Soares fetal liver spleen INFILS"
 /sex="male"
 /dev_stage="20 week-post conception fetus"
 /lab_host="DH10B (ampicillin resistant)"
 /note="Organ: Liver and Spleen; Vector: pTT73D (Pharmacia)
 with a modified polylinker; Site_1: Pac I; Site_2: Eco RI;
 1st strand cDNA was primed with a Pac I - oligo(dT) primer
 [5', AACTGAGAGATATTAAGATCTTTTCTTTTCTTTT 3'],
 double-stranded cDNA was ligated to Eco RI adaptors
 (Pharmacia), digested with Pac I and cloned into the Pac I
 and Eco RI sites of the modified pTT73 vector. Library
 went through one round of normalization. Library
 constructed by Bento Soares and M.Fatima Bonaldo."
 BASE COUNT 111 a 100 c 117 g 142 t 9 others
 ORIGIN
 Query Match 88.2%; Score 19.4; DB 14; Length 479;
 Best Local Similarity 95.2%; Pred. No. 1.8e+02;
 Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 GCTCTTTGGGACAGGAGCT 21
 |||||
 Db 103 GCTCTTTGGGACAGGAGCT 123
 |||||
 RESULT 5
 BF914566
 LOCUS IL3-UT0114-011200-362-E06 UT0114 Homo sapiens cDNA, mRNA sequence.
 DEFINITION BF914566
 ACCESSION BF914566
 VERSION BF914566.1 GI:12306024
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 851)
 AUTHORS NTH-MGC
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs@mail.nih.gov
 Tissue Procurement: ATCC/DCT/DTP
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Agencourt Bioscience Corporation

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 620)
 Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R.,
 Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.,
 Goldman, G.H., Carvalho, A.F., Matsukuma, A., Baia, G.S., Simpson, D.H.,
 Brunstein, A., de Oliveira, P.S., Bucher, P., Jorgensen, C.V., O'Hare
 W.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and
 Simpson, A.J.
 TITLE Shotgun sequencing of the human transcriptome with ORF expressed
 sequence tags
 JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
 MEDLINE 20202663
 COMMENT Contact: Simpson A.J.G.
 Laboratory of Cancer Genetics
 Ludwig Institute for Cancer Research
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
 Brazil
 Tel: +55-11-2704922
 Fax: +55-11-2707001
 Email: asimpson@ludwig.org.br
 This sequence was derived from the FAPESP/LICR Human Cancer Genome
 Project. This entry can be seen in the following URL
 (http://www.ludwig.org.br/scripts/gethtml2.pl?tl=IL3&t2=IL3-UT0114-
 011200-362-E06&t3=2000-12-01&t4=1)
 Seq primer: puc 18 forward
 High quality sequence stop: 619.
 Location/Qualifiers
 1..620
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_lib="UT0114"
 /dev_stage="Adult"
 /note="Organ: uterus tumor; Vector: puc18; Site 1: SmaI;
 Site 2: SmaI; A mini-library was made by cloning products
 derived from ORESTES PCR (U.S. Letters Patent application
 No. 196,716 - Ludwig Institute for Cancer Research)
 profiles into the pUC 18 vector. Reverse transcription of
 tissue mRNA and cDNA amplification were performed under
 low stringency conditions."
 BASE COUNT 181 a 157 c 150 g 132 t
 ORIGIN
 Query Match 88.2%; Score 19.4; DB 12; Length 620;
 Best Local Similarity 95.2%; Pred. No. 1.8e+02;
 Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 GCTCTTTGGGACAGGAGCT 21
 |||||
 Db 73 GCTCTTTGGGACAGGAGCT 93
 |||||
 RESULT 6
 BQ221568
 LOCUS AGNCOURT_7559207 NIH_MGC_72 Homo sapiens cDNA clone IMAGE:6046477
 DEFINITION 5', mRNA sequence.
 ACCESSION BQ221568
 VERSION BQ221568.1 GI:20402968
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 851)
 AUTHORS NTH-MGC
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs@mail.nih.gov
 Tissue Procurement: ATCC/DCT/DTP
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Agencourt Bioscience Corporation

Mon Apr 21 10:30:29 2003

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Plate: LLAM13292 row: f column: 14
 High quality sequence stop: 637.

FEATURES

source

1. .851
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:6046477"
 /clone_lib="NIH MGC 72"
 /tissue_type="melanotic melanoma"
 /lab_host="DH10B (phage-resistant)"
 /note="Orgân: skin; Vector: pCMV-SPORT6; Site 1: NotI; Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt. Average insert size 2 kb. Library constructed by Life Technologies."

BASE COUNT 229 a 234 c 231 g 157 t

ORIGIN

Query Match 88.2%; Score 19.4; DB 14; Length 851;
 Best Local Similarity 95.2%; Pred. No. 1.9e+02;
 Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GCTCTTTGGGACAGGAAGGT 21

Db 86 GCTCTTTGGGACAGGAAGCT 106

RESULT 7

BG481336

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

BG481336 874 bp mRNA linear EST 21-MAR-2001
 602528680F1 NIH_MGC_21 Homo sapiens cDNA clone IMAGE:4652395 5',
 mRNA sequence.

ACCESSION BG481336.1 GI:13413615

VERSION EST.

KEYWORDS human.

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 874)
 NIH-MGC <http://mgc.nci.nih.gov/>.
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: ATCC

cDNA Library Preparation: Ling Hong/Rubin Laboratory
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 Plate: LLCM1436 row: o column: 20
 High quality sequence stop: 688.

BASE COUNT 240 a 230 c 250 g 154 t

FEATURES

source

1. .874
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:4652395"
 /clone_lib="NIH MGC 21"
 /tissue_type="choriocarcinoma"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: placenta; Vector: pOTB7; Site 1: XhoI; Site 2: EcoRI; cDNA made by oligo-dt priming.
 Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Size-selected for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies)."

BASE COUNT 240 a 230 c 250 g 154 t

ORIGIN

Query Match 88.2%; Score 19.4; DB 12; Length 874;
 Best Local Similarity 95.2%; Pred. No. 1.9e+02;
 Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GCTCTTTGGGACAGGAAGGT 21

Db 116 GCTCTTTGGGACAGGAAGCT 136

RESULT 8

BG748346

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

BG748346 889 bp mRNA linear EST 15-MAY-2001
 602705902F1 NIH_MGC_43 Homo sapiens cDNA clone IMAGE:4842553 5',
 mRNA sequence.

ACCESSION BG748346

VERSION BG748346.1 GI:14058999

KEYWORDS EST.

SOURCE human.

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 889)
 NIH-MGC <http://mgc.nci.nih.gov/>.
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: ATCC

cDNA Library Preparation: Ling Hong/Rubin Laboratory
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 Plate: LLCM1677 row: c column: 02
 High quality sequence stop: 773.

BASE COUNT 242 a 241 c 235 g 171 t

FEATURES

source

1. .889
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:4842553"
 /clone_lib="NIH MGC 43"
 /tissue_type="normal pigmented retinal epithelium"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: eye; Vector: pOTB7; Site 1: XhoI; Site 2: EcoRI; cDNA made by oligo-dt priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies)."

BASE COUNT 242 a 241 c 235 g 171 t

ORIGIN

Query Match 88.2%; Score 19.4; DB 12; Length 889;
 Best Local Similarity 95.2%; Pred. No. 1.9e+02;
 Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GCTCTTTGGGACAGGAAGGT 21

Db 62 GCTCTTTGGGACAGGAAGCT 82

RESULT 9

B1871125/c

LOCUS

DEFINITION

ACCESSION

VERSION

B1871125 264 bp mRNA linear EST 11-OCT-2001
 603395008F1 NIH_MGC_90 Homo sapiens cDNA clone IMAGE:5405040 5',
 mRNA sequence.

ACCESSION B1871125

VERSION B1871125.1 GI:16044798

KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
1 (bases 1 to 264) Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

AUTHORS NIH-MGC <http://mgs.nci.nih.gov/>.

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished (1999)

COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs@mail.nih.gov
Tissue procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLAM12032 row: p column: 01
High quality sequence start: 9
High quality sequence stop: 264.

FEATURES
source
Location/Qualifiers
1..264
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5405040"
/clone_lib="NIH_MGC_90"
/tissue_type="adenocarcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: liver; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally; oligo-dT primed.
Average insert size 1.7 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: This is a NIH_MGC Library."

BASE COUNT 87 a 82 c 48 g 47 t

ORIGIN

Query Match 85.5%; Score 18.8; DB 13; Length 264;
Best Local Similarity 90.9%; Pred. No. 3.2e+02;
Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GCTCTTTGGGACAGGAAGTC 22
|||||
Db 223 GCTCTGTGGGGCAGGAAGTC 202
|||||

RESULT 10
BM484358/c
LOCUS 538191 MARC 2P1G Sus scrofa cDNA 5', mRNA linear EST 05-FEB-2002

DEFINITION BM484358

ACCESSION BM484358

VERSION BM484358.1 GI:18534686

KEYWORDS EST.

SOURCE pig.

ORGANISM Sus scrofa

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
1 (bases 1 to 309) Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.

AUTHORS Fahrenkrug, S.C., Freking, B.A., Rohrer, G.A., Smith, T.P.L., Casas, E.,
Stone, R.T., Heaton, M.P., Grosse, W.M., Bennett, G.A., Laegreid, W.W.,
and Keele, J.W.

TITLE Design and use of two pooled tissue normalized cDNA libraries for

JOURNAL EST discovery in swine

COMMENT Unpublished (2000)
Contact: Smith TPL
USDA, ARS, US Meat Animal Research Center
PO Box 166, Clay Center, NE 68933-0166, USA
Tel: 402 762 4366
Fax: 402 762 4390
Email: smith@mail.marc.usda.gov
Single pass sequencing. Bases called and alt trimmed with phred
v0.980904.e. Vector identified by cross_match with the -minscore 18
and -minmatch 12 options.

BASE COUNT 87 a 82 c 48 g 47 t

ORIGIN

Query Match 85.5%; Score 18.8; DB 13; Length 264;
Best Local Similarity 90.9%; Pred. No. 3.2e+02;
Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GCTCTTTGGGACAGGAAGTC 22
|||||
Db 223 GCTCTGTGGGGCAGGAAGTC 202
|||||

RESULT 10
BM484358/c
LOCUS 538191 MARC 2P1G Sus scrofa cDNA 5', mRNA linear EST 05-FEB-2002

DEFINITION BM484358

ACCESSION BM484358

VERSION BM484358.1 GI:18534686

KEYWORDS EST.

SOURCE pig.

ORGANISM Sus scrofa

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
1 (bases 1 to 309) Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.

AUTHORS Fahrenkrug, S.C., Freking, B.A., Rohrer, G.A., Smith, T.P.L., Casas, E.,
Stone, R.T., Heaton, M.P., Grosse, W.M., Bennett, G.A., Laegreid, W.W.,
and Keele, J.W.

TITLE Design and use of two pooled tissue normalized cDNA libraries for

JOURNAL EST discovery in swine

COMMENT Unpublished (2000)
Contact: Smith TPL
USDA, ARS, US Meat Animal Research Center
PO Box 166, Clay Center, NE 68933-0166, USA
Tel: 402 762 4366
Fax: 402 762 4390
Email: smith@mail.marc.usda.gov
Single pass sequencing. Bases called and alt trimmed with phred
v0.980904.e. Vector identified by cross_match with the -minscore 18
and -minmatch 12 options.

PCR Primers
FORWARD: AGGAAACAGCTATGACCAT
BACKWARD: GTTTCCTCCAGTCACGACG
Plate: 3 row: M column: 20
Seq primer: ATTAGTGACACTATAG.

FEATURES
source
Location/Qualifiers
1..309
/organism="Sus scrofa"
/db_xref="taxon:9823"
/clone_lib="MARC 2P1G"
/tissue_type="pooled"
/lab_host="DH10B"
/note="Vector: pCMV SPORT6; Site 1: NotI; Site 2: SalI;
Library made from pooled tissue from testis, ovary,
endometrium, hypothalamus, pituitary, and placenta."

BASE COUNT 72 a 93 c 86 g 58 t

ORIGIN

Query Match 85.5%; Score 18.8; DB 13; Length 309;
Best Local Similarity 90.9%; Pred. No. 3.2e+02;
Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GCTCTTTGGGACAGGAAGTC 22
|||||
Db 184 GCTCTTTGAGGACTGGAAGTC 163
|||||

RESULT 11
BF189397/c
LOCUS 234957 MARC 2P1G Sus scrofa cDNA 5', mRNA linear EST 02-NOV-2000

DEFINITION BF189397

ACCESSION BF189397

VERSION BF189397.1 GI:11072766

KEYWORDS EST.

SOURCE pig.

ORGANISM Sus scrofa

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
1 (bases 1 to 373) Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.

AUTHORS Fahrenkrug, S.C., Freking, B.A., Rohrer, G.A., Smith, T.P.L., Casas, E.,
Stone, R.T., Heaton, M.P., Grosse, W.M., Bennett, G.A., Laegreid, W.W.,
and Keele, J.W.

TITLE Design and use of two pooled tissue normalized cDNA libraries for

JOURNAL EST discovery in swine

COMMENT Unpublished (2000)
Contact: Smith TPL
USDA, ARS, US Meat Animal Research Center
PO Box 166, Clay Center, NE 68933-0166, USA
Tel: 402 762 4366
Fax: 402 762 4390
Email: smith@mail.marc.usda.gov
Single pass sequencing. Bases called and alt trimmed with phred
v0.980904.e. Vector identified by cross_match with the -minscore 18
and -minmatch 12 options.

BASE COUNT 91 a 100 c 106 g 76 t

ORIGIN

Query Match 85.5%; Score 18.8; DB 12; Length 373;

FEATURES
source
Location/Qualifiers
1..373
/organism="Sus scrofa"
/db_xref="taxon:9823"
/clone_lib="MARC 2P1G"
/tissue_type="pooled"
/lab_host="DH10B"
/note="Vector: pCMV SPORT6; Site 1: NotI; Site 2: SalI;
Library made from pooled tissue from testis, ovary,
endometrium, hypothalamus, pituitary, and placenta."

Mon Apr 21 10:30:29 2003

```

REFERENCE
AUTHORS      Zhao,S., Shetty,J., Shatsman,S., Tsegaye,G., Geer,K., Shvartsbeyn
              A., Gebregorgis,E., Overton,L., Russell,D., Chen,D., Riggs,F., de
              Jong,P. and Fraser,C.M.
TITLE        Rat BAC End Sequences from Library CHORI-230 EcoRI segment
JOURNAL      Unpublished (1999)
COMMENT      Other GSSs: CH230-185B17.TV
              Contact: Shaying Zhao
              Department of Eukaryotic Genomics
              The Institute for Genomic Research
              9712 Medical Center Dr., Rockville, MD 20850, USA
              Tel: 301 838 0200
              Fax: 301 838 0208
              Email: szhao@tigr.org
              Clones are derived from the rat BAC library CHORI-230
              (http://www.chori.org/bacpac/rat230.htm). For BAC library
              availability, please contact Pieter de Jong (pdejong@mail.cho.org).
              Clones may be purchased from BACPAC Resources
              (http://www.chori.org/bacpac/or ering information.htm). BAC end
              page: http://www.tigr.org/tldb/bac_ends/rat/bac_end_intro.html
              Plate: 185 row: B column: 17
              Seq primer: SP6
              Class: BAC ends
              Location/Qualifiers
                  1..566
                     /organism="Rattus norvegicus"
                     /strain="BN/SnHsd/MCW"
                     /db_xref="taxon:10116"
                     /clone="CH230-185B17"
                     /clone_lib="CHORI-230 Segment 1"
                     /sex="Female"
                     /cell_type="Brain"
                     /note="vector: pTARAC2.1; Site 1: EcoRI; Site 2: EcoRI;
                     CHORI-230 Rat (BN/SnHsd/MCW) BAC library produced by
                     Pieter de Jong"
BASE COUNT   211 a 121 c 103 g 131 t
ORIGIN
Query Match  83.6%; Score 18.4; DB 17; Length 566;
Best Local Similarity 95.0%; Pred. No. 5.1e+02;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 TCCTTGGGACAGGAGGTC 22
Db 484 TCCTTGGGACAGGAGGTC 465

RESULT 14
LOCUS      BH098064
DEFINITION RPI-24-271L15.TJ RPI-24 Mus musculus genomic clone RPI-24-271L15
            , DNA sequence.
ACCESSION  BH098064
VERSION     BH098064.1 GI:14919489
KEYWORDS   GSS.
SOURCE     house mouse.
ORGANISM   Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE  1 (bases 1 to 680)
AUTHORS    Zhao,S., Geer,K., Malek,J., Shatsman,S., Akinret,B., Levins,M.,
            Tsegaye,G., Overton,L., Russell,D., Chen,D., Riggs,F., de
            Jong,P. and Fraser,C.M.
            Russell,D., de Jong,P. and Fraser,C.M.
            Mouse BAC End Sequences from Library RPI-24
            Unpublished (1999)
            Other GSSs: RPI-24-271L15.TV
            Contact: Shaying Zhao
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: szhao@tigr.org

TITLE      Unpublished (1999)
JOURNAL    RPI-24-271L15.TV
COMMENT    Contact: Shaying Zhao
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: szhao@tigr.org
    
```

```

REFERENCE
AUTHORS      Zhao,S., Shetty,J., Shatsman,S., Tsegaye,G., Geer,K., Shvartsbeyn
              A., Gebregorgis,E., Overton,L., Russell,D., Chen,D., Riggs,F., de
              Jong,P. and Fraser,C.M.
TITLE        Rat BAC End Sequences from Library CHORI-230 EcoRI segment
JOURNAL      Unpublished (1999)
COMMENT      Other GSSs: CH230-185B17.TV
              Contact: Shaying Zhao
              Department of Eukaryotic Genomics
              The Institute for Genomic Research
              9712 Medical Center Dr., Rockville, MD 20850, USA
              Tel: 301 838 0200
              Fax: 301 838 0208
              Email: szhao@tigr.org
              Clones are derived from the rat BAC library CHORI-230
              (http://www.chori.org/bacpac/rat230.htm). For BAC library
              availability, please contact Pieter de Jong (pdejong@mail.cho.org).
              Clones may be purchased from BACPAC Resources
              (http://www.chori.org/bacpac/or ering information.htm). BAC end
              page: http://www.tigr.org/tldb/bac_ends/rat/bac_end_intro.html
              Plate: 185 row: B column: 17
              Seq primer: SP6
              Class: BAC ends
              Location/Qualifiers
                  1..566
                     /organism="Rattus norvegicus"
                     /strain="BN/SnHsd/MCW"
                     /db_xref="taxon:10116"
                     /clone="CH230-185B17"
                     /clone_lib="CHORI-230 Segment 1"
                     /sex="Female"
                     /cell_type="Brain"
                     /note="vector: pTARAC2.1; Site 1: EcoRI; Site 2: EcoRI;
                     CHORI-230 Rat (BN/SnHsd/MCW) BAC library produced by
                     Pieter de Jong"
BASE COUNT   211 a 121 c 103 g 131 t
ORIGIN
Query Match  83.6%; Score 18.4; DB 17; Length 566;
Best Local Similarity 95.0%; Pred. No. 5.1e+02;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 TCCTTGGGACAGGAGGTC 22
Db 484 TCCTTGGGACAGGAGGTC 465

RESULT 14
LOCUS      BH098064
DEFINITION RPI-24-271L15.TJ RPI-24 Mus musculus genomic clone RPI-24-271L15
            , DNA sequence.
ACCESSION  BH098064
VERSION     BH098064.1 GI:14919489
KEYWORDS   GSS.
SOURCE     house mouse.
ORGANISM   Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE  1 (bases 1 to 680)
AUTHORS    Zhao,S., Geer,K., Malek,J., Shatsman,S., Akinret,B., Levins,M.,
            Tsegaye,G., Overton,L., Russell,D., Chen,D., Riggs,F., de
            Jong,P. and Fraser,C.M.
            Russell,D., de Jong,P. and Fraser,C.M.
            Mouse BAC End Sequences from Library RPI-24
            Unpublished (1999)
            Other GSSs: RPI-24-271L15.TV
            Contact: Shaying Zhao
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: szhao@tigr.org

TITLE      Unpublished (1999)
JOURNAL    RPI-24-271L15.TV
COMMENT    Contact: Shaying Zhao
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: szhao@tigr.org
    
```

```

REFERENCE
AUTHORS      Zhao,S., Shetty,J., Shatsman,S., Tsegaye,G., Geer,K., Shvartsbeyn
              A., Gebregorgis,E., Overton,L., Russell,D., Chen,D., Riggs,F., de
              Jong,P. and Fraser,C.M.
TITLE        Rat BAC End Sequences from Library CHORI-230 EcoRI segment
JOURNAL      Unpublished (1999)
COMMENT      Other GSSs: CH230-185B17.TV
              Contact: Shaying Zhao
              Department of Eukaryotic Genomics
              The Institute for Genomic Research
              9712 Medical Center Dr., Rockville, MD 20850, USA
              Tel: 301 838 0200
              Fax: 301 838 0208
              Email: szhao@tigr.org
              Clones are derived from the rat BAC library CHORI-230
              (http://www.chori.org/bacpac/rat230.htm). For BAC library
              availability, please contact Pieter de Jong (pdejong@mail.cho.org).
              Clones may be purchased from BACPAC Resources
              (http://www.chori.org/bacpac/or ering information.htm). BAC end
              page: http://www.tigr.org/tldb/bac_ends/rat/bac_end_intro.html
              Plate: 185 row: B column: 17
              Seq primer: SP6
              Class: BAC ends
              Location/Qualifiers
                  1..566
                     /organism="Rattus norvegicus"
                     /strain="BN/SnHsd/MCW"
                     /db_xref="taxon:10116"
                     /clone="CH230-185B17"
                     /clone_lib="CHORI-230 Segment 1"
                     /sex="Female"
                     /cell_type="Brain"
                     /note="vector: pTARAC2.1; Site 1: EcoRI; Site 2: EcoRI;
                     CHORI-230 Rat (BN/SnHsd/MCW) BAC library produced by
                     Pieter de Jong"
BASE COUNT   211 a 121 c 103 g 131 t
ORIGIN
Query Match  83.6%; Score 18.4; DB 17; Length 566;
Best Local Similarity 95.0%; Pred. No. 5.1e+02;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 TCCTTGGGACAGGAGGTC 22
Db 484 TCCTTGGGACAGGAGGTC 465

RESULT 14
LOCUS      BH098064
DEFINITION RPI-24-271L15.TJ RPI-24 Mus musculus genomic clone RPI-24-271L15
            , DNA sequence.
ACCESSION  BH098064
VERSION     BH098064.1 GI:14919489
KEYWORDS   GSS.
SOURCE     house mouse.
ORGANISM   Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE  1 (bases 1 to 680)
AUTHORS    Zhao,S., Geer,K., Malek,J., Shatsman,S., Akinret,B., Levins,M.,
            Tsegaye,G., Overton,L., Russell,D., Chen,D., Riggs,F., de
            Jong,P. and Fraser,C.M.
            Russell,D., de Jong,P. and Fraser,C.M.
            Mouse BAC End Sequences from Library RPI-24
            Unpublished (1999)
            Other GSSs: RPI-24-271L15.TV
            Contact: Shaying Zhao
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: szhao@tigr.org

TITLE      Unpublished (1999)
JOURNAL    RPI-24-271L15.TV
COMMENT    Contact: Shaying Zhao
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: szhao@tigr.org
    
```

Clones are derived from the mouse BAC library RPCI-24. For BAC library availability, please contact Pieter de Jong (pdejong@mail.cho.org). Clones may be purchased from BACPAC Resources (<http://www.choi.org/bacpac/orderingframe.html>). BAC end page: http://www.tigr.org/tcb/bac_ends/mouse/bac_end_intro.html. Plate: 271 row: L column: 15
Seq primer: Sp6
Class: BAC ends.

FEATURES

source
Location/Qualifiers
1. .680
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="RPCI-24-271L15"
/clone_lib="RPCI-24"
/sex="Male"
/cell_type="Spleen/Brain"
/note="Vector: pTARBAC1; Site 1: BamHI; Site 2: BamHI; RPCI-24 Mouse BAC library produced by Pieter de Jong. The library was cloned in the pTARBAC1 cloning vector at the BamHI sites using MboI partially digested male C57BL/6J DNA."
BASE COUNT 169 a 155 c 159 g 197 t
ORIGIN

Query Match 83.6%; Score 18.4; DB 17; Length 680;
Best Local Similarity 95.0%; Pred. No. 5.1e+02;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 2 CTCTTTGGGACAGGAAGT 21
|||||
Db 67 CTCTTTGGGACAGGAAGT 86
|||||

RESULT 15
BG480515
LOCUS
DEFINITION 602529520F1 NIH_MGC_21 Homo sapiens cDNA clone IMAGE:4652981 5',
mRNA sequence.
ACCESSION BG480515
VERSION BG480515.1 GI:13412705
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 856)
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone Distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Plate: LLCW1438 row: h column: 06
High quality sequence stop: 661.
Location/Qualifiers
1. .856
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4652981"
/clone_lib="NIH_MGC_21"
/tissue_type="Choriocarcinoma"
/lab_host="DH10B (phage-resistant)"
/note="Organ: placenta; Vector: pOTB7; Site 1: XhoI; Site 2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGACGAG(G). Size-selected >500bp

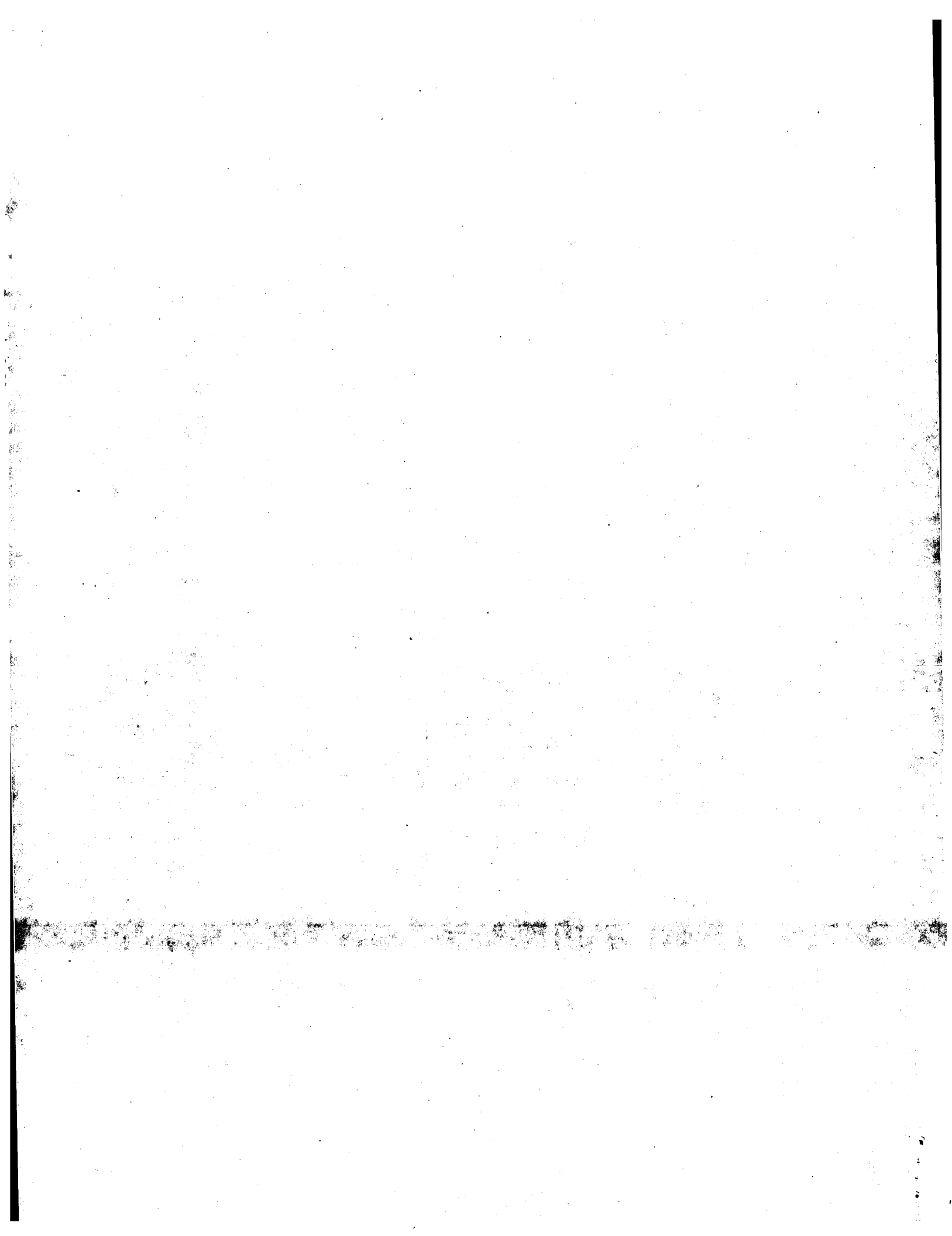
FEATURES

source
Location/Qualifiers
1. .856
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4652981"
/clone_lib="NIH_MGC_21"
/tissue_type="Choriocarcinoma"
/lab_host="DH10B (phage-resistant)"
/note="Organ: placenta; Vector: pOTB7; Site 1: XhoI; Site 2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGACGAG(G). Size-selected >500bp

BASE COUNT 225 a 250 c 233 g 148 t
ORIGIN
Query Match 83.6%; Score 18.4; DB 12; Length 856;
Best Local Similarity 95.0%; Pred. No. 5.2e+02;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1 GCTCTTTGGGACAGGAAG 20
|||||
Db 117 GCTCTTTGGGACAGGAATG 136
|||||

Search completed: April 18, 2003, 07:32:01
Job time : 770.167 secs

for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies)."



GenCore version 5.1.4 p5 4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 18, 2003, 04:46:26 ; Search time 292 seconds
(without alignments)
2192.677 Million cell updates/sec

Title: US-09-270-437D-14
Perfect score: 22
Sequence: 1 gacgttgacacacgcgggtttct 22

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl.*
1: gb.ba.*
2: gb.htg.*
3: gb.in.*
4: gb.om.*
5: gb.ov.*
6: gb.pat.*
7: gb.ph.*
8: gb.pl.*
9: gb.pr.*
10: gb.ro.*
11: gb.sts.*
12: gb.ey.*
13: gb.un.*
14: gb.vi.*
15: em.ba.*
16: em.fun.*
17: em.hum.*
18: em.in.*
19: em.mu.*
20: em.om.*
21: em.or.*
22: em.ov.*
23: em.pat.*
24: em.ph.*
25: em.pl.*
26: em.ro.*
27: em.sts.*
28: em.un.*
29: em.vi.*
30: em.htg.hum.*
31: em.htg.inv.*
32: em.htg.other.*
33: em.htg.mus.*
34: em.htg.pln.*
35: em.htg.rod.*
36: em.htg.nam.*
37: em.htg.vrt.*
38: em.sy.*
39: em.htgo.hum.*
40: em.htgo.mus.*
41: em.htgo.other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
C 1	22	100.0	2010	9	AF117107	AF117107 Homo sapi
C 2	22	100.0	3283	6	AR171857	AR171857 Sequence
C 3	22	100.0	3412	6	AR171865	AR171865 Sequence
C 4	22	100.0	3642	9	BC021290	BC021290 Homo sapi
C 5	22	100.0	3667	9	AF057352	AF057352 Homo sapi
C 6	22	100.0	50320	2	AC126373	AC126373 Homo sapi
C 7	22	100.0	200947	2	AC016961	AC016961 Homo sapi
C 8	20.4	92.7	124849	2	AC094924	AC094924 Homo sapi
C 9	20.4	92.7	161691	2	AC126477	AC126477 Rattus no
C 10	18.8	85.5	91084	9	AL596177	AL596177 Human DNA
C 11	18.8	85.5	112259	2	AC026579	AC026579 Homo sapi
C 12	18.8	85.5	142971	9	AC020629	AC020629 Homo sapi
C 13	18.8	85.5	174623	2	AC108224	AC108224 Oryza sat
C 14	18.8	85.5	182695	2	AC015706	AC015706 Homo sapi
C 15	18.4	83.6	2219	1	NGU43735	U43735 Neisseria g
C 16	18.4	83.6	11099	1	AE002425	AE002425 Neisseria
C 17	18.4	83.6	11579	1	AE002453	AE002453 Neisseria
C 18	18.4	83.6	311321	1	NMA322491	AL162754 Neisseria
C 19	18.4	83.6	329861	1	NMA522491	AL162756 Neisseria
C 20	18.4	83.6	349980	6	AX044030	AX044030 Sequence
C 21	18.4	83.6	349980	6	AX044031	AX044031 Sequence
C 22	17.8	80.9	178212	2	AC107200	AC107200 Rattus no
C 23	17.4	79.1	4029	1	NMA277988	AJ277988 Natrialba
C 24	17.2	78.2	1929	14	AF121457	AF121457 Hyphantri
C 25	17.2	78.2	2341	1	RERORF1188	D30033 Rhodococcus
C 26	17.2	78.2	2809	8	TAU51304	U51304 Triticum ae
C 27	17.2	78.2	7001	1	AB016078	AB016078 Rhodococ
C 28	17.2	78.2	9515	1	AE007083	AE007083 Mycobacte
C 29	17.2	78.2	12948	1	AE006162	AE006162 Pasteurel
C 30	17.2	78.2	27548	1	MTCY27	295208 Mycobacteri
C 31	17.2	78.2	55380	8	NC21D9	AL807373 Neurospor
C 32	17.2	78.2	56510	2	AC020215	AC020215 Drosophil
C 33	17.2	78.2	62670	2	AC010566	AC010566 Drosophil
C 34	17.2	78.2	67375	2	AC104767	AC104767 Homo sapi
C 35	17.2	78.2	67375	2	AC104767	AC104767 Homo sapi
C 36	17.2	78.2	79554	8	NCB11B23	AC104767 Homo sapi
C 37	17.2	78.2	113201	9	AP004290	AL669991 Neurospor
C 38	17.2	78.2	120524	9	AC008443	AP004290 Homo sapi
C 39	17.2	78.2	125798	9	AL390248	AC008443 Homo sapi
C 40	17.2	78.2	148824	2	AC025640	AL390248 Human DNA
C 41	17.2	78.2	151343	8	AC021893	AC025640 Homo sapi
C 42	17.2	78.2	159462	2	AC120991	AC021893 Genomic S
C 43	17.2	78.2	160930	9	AC010093	AC120991 Oryza sat
C 44	17.2	78.2	183455	2	AC008620	AC010093 Homo sapi
C 45	17.2	78.2	187226	9	AC104980	AC008620 Homo sapi
C 45	17.2	78.2	187226	9	AC104980	AC104980 Homo sapi

ALIGNMENTS

RESULT 1
AF117107/c
LOCUS AF117107
DEFINITION Homo sapiens IGF-II mRNA-binding protein 2 (IMP-2) mRNA, complete cds.
ACCESSION AF117107
VERSION AF117107.1
KEYWORDS GI:4191609
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 2010)
REFERENCE
AUTHORS Nielsen J., Christiansen J., Lykke-Andersen J., Johnsen A.H., Wewer U.M. and Nielsen F.C.
2010 bp mRNA linear PRI 26-JAN-1999
Homo sapiens IGF-II mRNA-binding protein 2 (IMP-2) mRNA, complete

us-09-270-437d-14.rge

Mon Apr 21 10:30:32 2003

```

Query Match      100.0%; Score 22; DB 6; Length 3283;
Best Local Similarity 100.0%; Pred. No. 9.6;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAGCTTGACAAACGCGGTTTCT 22
    |||||
Db 444 GAGCTTGACAAACGCGGTTTCT 423

RESULT 3
LOCUS AR171865/c 3412 bp DNA linear PAT 17-DEC-2001
DEFINITION Sequence 6 from patent US 6297364.
ACCESSION AR171865
VERSION AR171865.1 GI:17910815
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 3412)
AUTHORS Chen,Y.-T., Gure,A., Tsang,S., Stockert,E., Jager,E., Alexander,K.
and Old,L.J.
TITLE Isolated nucleic acid molecule encoding cancer associated antigen,
the antigen itself, and uses thereof
JOURNAL Patent: US 6297364-A 6 02-OCT-2001;
FEATURES Location/Qualifiers
source 1. 3412
/organism="unknown"
BASE COUNT 970 a 887 c 804 g 750 t 1 others
ORIGIN

Query Match      100.0%; Score 22; DB 6; Length 3412;
Best Local Similarity 100.0%; Pred. No. 9.6;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAGCTTGACAAACGCGGTTTCT 22
    |||||
Db 444 GAGCTTGACAAACGCGGTTTCT 423

RESULT 4
LOCUS BC021290/c 3642 bp mRNA linear PRI 22-JAN-2002
DEFINITION Homo sapiens, IGF-II mRNA-binding protein 2, clone MGC:29539
IMAGE:5090334, mRNA, complete cds.
ACCESSION BC021290
VERSION BC021290.1 GI:18204200
KEYWORDS MGC.
SOURCE Homo sapiens.
ORGANISM Homo sapiens.
REFERENCE 1 (bases 1 to 3642)
AUTHORS Strausberg,R.
TITLE Direct Submission
JOURNAL Submitted (14-JAN-2002) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
REMARK NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT Contact: MGC help desk
Email: ggapbs-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: National Institutes of Health Intramural
Sequencing Center (NISC),
Gaithersburg, Maryland;
Web site: http://www.nisc.nih.gov/
Contact: nisc.mgc@nih.gov
Shevchenko,Y., Wetherby,K.D., Beckstrom-Sternberg,S.M.,
Benjamin,B., Blakesley,R.W., Bouffard,G.G., Brinkley,C., Brooks,S.,
Dietrich,N.L., Guan,X., Gupta,J., Ho,S.-L., Karlins,E., Legaspi,R.,

```

Lim, M., Maduro, Q.L., Mastello, C., Mastrian, S.D., McCloskey, J.C., McDowell, J., Pearson, R., Snyder, B., Stantripop, S., Thomas, P.J., Tjongson, E.E., Touchman, J.W., Tourgeon, C., Vogt, J.L., Walker, M.A., Zhang, L.-H. and Green, B.D.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LINL at: <http://image.llnl.gov> Series: IRAL Plate: 39 Row: 0 Column: 21 This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 5729883.

FEATURES
source
1. .3642
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="MGC:29539 IMAGE:5090334"
/tissue type="Pancreas, epithelioid carcinoma"
/clone_lib="NIH MGC_42"
/lab_host="DH10B-R"
/notes="Vector: pOTB7"
77. .1873
/codon_start=1
/product="IGF-II mRNA-binding protein 2"
/protein_id="AAH21290.1"
/db_xref="GI:18204201"
/translation="MNKLYIGNLSPAVTADDLRLQFGDKLPLAGQVLLKSGYAFVDY
PQONWAIETLSGKVLHGKIMEVDYSVSKLRSKIQIRNPPHLOWEVLGLLA
OYGTVENVEQNTDTETAVVNTVATREAEKIAEKLSGHOFENYSFKISYIPDEEVS
SPSPQRAQRGDHSREGCHAPGTSQARQIDFPLRLVPTQFVGAIGKEGLTIKNI
TKQTSRDVIRKENSAGAEKPVTHATPEGTSEACRMILEIMQKAEDETKLAEIP
KLILAHGVLGRILGKGRNLKKIETHETGKTIITSSQLDLSIYNPRTITVKGTVEACA
SAEIMKLEAFENDMLAVNQANLIPGLNLSALGIFSTGLSVLSPAPRGPAGPA
APYHFTTSGTSSLYPHHGFPPPHHSYEQEIVNLFTPTQAVGAIIGKGAHTK
QARFAGASIKIAPAGPDVSRMVIITGPPEAQFKAGRIFGKLEENFNFKEVK
LEAHIRVPSSTAGRVGKGTVNELNLTSAEIVVPRDQTPDENEIVVRIIGHFPA
SQTAKRIREIVQVQKQEQKYPQGVASQSK"

CDS
BASE COUNT 1077 a 919 c 849 g 737 t
ORIGIN
Query Match 100.0%; Score 22; DB 9; Length 3642;
Best Local Similarity 100.0%; Pred. No. 9.5;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GACGTTGACACGGCGGTTTCT 22
|||||
Db 448 GACGTTGACACGGCGGTTTCT 427

RESULT 5
AF057352/c
LOCUS AF057352 Homo sapiens hepatocellular carcinoma autoantigen (p62) mRNA, complete cds.
DEFINITION AF057352 3667 bp mRNA linear PRI 23-MAY-1999
ACCESSION AF057352.1 GI:4883680
VERSION AF057352.1
KEYWORDS Homo sapiens.
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 3667)
Zhang, J.Y., Chan, E.K., Peng, X.X. and Tan, E.M.
A novel cytoplasmic protein with RNA-binding motifs is an autoantigen in human hepatocellular carcinoma
J. Exp. Med. 189 (7), 1101-1110 (1999)
JOURNAL 99207072
MEDLINE 10190901
PUBMED 10190901
REFERENCE 2 (bases 1 to 3667)
Zhang, J.Y., Chan, E.K.L., Peng, X.X. and Tan, E.M.
Direct Submission
Submitted (03-APR-1998) Molecular & Experimental Medicine, The Scripps Research Institute, 10550 N. Torrey Pines Road, La Jolla, CA 92037, USA

FEATURES
source
1. .3667
/organism="Homo sapiens"
/db_xref="taxon:9606"
gene
1. .3667
/gene="p62"
436. .2106
/genes="p62"
/notes="RNA-binding protein"
/codon_start=1
/product="hepatocellular carcinoma autoantigen"
/protein_id="AAD31596.1"
/db_xref="GI:4883681"
/translation="MNKLYIGNLSPAVTADDLRLQFGDKLPLAGQVLLKSGYAFVDY
PQONWAIETLSGKVLHGKIMEVDYSVSKLRSKIQIRNPPHLOWEVLGLLA
OYGTVENVEQNTDTETAVVNTVATREAEKIAEKLSGHOFENYSFKISYIPDEEVS
SPSPQRAQRGDHSREGCHAPGTSQARQIDFPLRLVPTQFVGAIGKEGLTIKNI
TKQTSRDVIRKENSAGAEKPVTHATPEGTSEACRMILEIMQKAEDETKLAEIP
KLILAHGVLGRILGKGRNLKKIETHETGKTIITSSQLDLSIYNPRTITVKGTVEACA
SAEIMKLEAFENDMLAVNTHSGYFSSLYPHHGFPPPHHSYEQEIVNLFTPTQAVGAIIGKGAHTK
QARFAGASIKIAPAGPDVSRMVIITGPPEAQFKAGRIFGKLEENFNFKEVK
GLKLEENFNFKEVKLEAHIRVPSSTAGRVGKGTVNELNLTSAEIVVPRDQTPDENEIVVRIIGHFPA
SQTAKRIREIVQVQKQEQKYPQGVASQSK"

BASE COUNT 1081 a 882 c 846 g 858 t
ORIGIN
Query Match 100.0%; Score 22; DB 9; Length 3667;
Best Local Similarity 100.0%; Pred. No. 9.5;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GACGTTGACACGGCGGTTTCT 22
|||||
Db 810 GACGTTGACACGGCGGTTTCT 789

RESULT 6
AC126373/c
LOCUS AC126373 Homo sapiens chromosome 18 clone CTD-2307113 map 18, LOW-PASS SEQUENCE SAMPLING.
DEFINITION AC126373
ACCESSION AC126373.1 GI:21699324
VERSION AC126373.1
KEYWORDS HTG; HTGS_PHASE0.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 50320)
Birren, B., Nussbaum, C. and Lander, E.
Homo sapiens chromosome 18, clone CTD-2307113
Unpublished
2 (bases 1 to 50320)
Birren, B., Nussbaum, C., Lander, E., Ali, A., Allen, N., Anderson, S., Barna, N., Bastien, V., Bloom, T., Boguslavskiy, L., Boukhgaltier, B., Camarata, J., Chang, J., Chazaro, B., Choepel, Y., Collymore, A., Cook, A., Cooke, P., DeArellano, K., Dewar, K., Diaz, J.S., Dodge, S., Faro, S., Ferreira, P., Fitzgerald, M., Gage, D., Gaigian, J., Gardyna, S., Gord, S., Graham, L., Grand-Pierre, N., Hagos, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karakas, A., Kells, C., Landers, T., Levine, R., Lindblad-Toh, K., Liu, G., MacLean, C., Macdonald, P., Major, J., Matthews, C., McCarthy, M., Meldrim, J., Meneus, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neil, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Raymond, C., Retta, R., Rise, C., Rogov, P., Roman, J., Roy, A., Schauer, S., Schupbach, R., Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.
Direct Submission
Submitted (05-JUL-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA

Mon Apr 21 10:30:32 2003

COMMENT

All repeats were identified using RepeatMasker:
 Smit, A.F.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>
 ----- Genome Center
 Center: Whitehead Institute/ MIT Center for Genome Research
 Center code: WIBR
 Web site: <http://www-seq.wi.mit.edu>
 Contact: sequence_submissions@genome.wi.mit.edu
 ----- Project Information
 Center project name: L27661
 Center clone name: 2307_1_13

NOTE: This record contains 59 individual
 * sequencing reads that have not been assembled into
 * contigs. Runs of N are used to separate the reads
 * and the order in which they appear is completely
 * arbitrary. Low-pass sequence sampling is useful for
 * identifying clones that may be gene-rich and allows
 * overlap relationships among clones to be deduced.
 * However, it should not be assumed that this clone
 * will be sequenced to completion. In the event that
 * the record is updated, the accession number will
 * be preserved.

1
 * 767 866: gap of 100 bp in length
 * 867 1574: contig of 708 bp in length
 * 1575 1674: gap of 100 bp
 * 1675 2412: contig of 738 bp in length
 * 2413 2512: gap of 100 bp
 * 2513 3283: contig of 771 bp in length
 * 3284 3383: gap of 100 bp
 * 3384 4124: contig of 741 bp in length
 * 4125 4224: gap of 100 bp
 * 4225 4938: contig of 714 bp in length
 * 4939 5038: gap of 100 bp
 * 5039 5920: contig of 782 bp in length
 * 5921 6685: contig of 765 bp in length
 * 6686 6785: gap of 100 bp
 * 6786 7580: contig of 795 bp in length
 * 7581 7680: gap of 100 bp
 * 7681 8432: contig of 752 bp in length
 * 8433 8532: gap of 100 bp
 * 8533 9281: contig of 749 bp in length
 * 9282 9381: gap of 100 bp
 * 9382 10136: contig of 755 bp in length
 * 10137 10236: gap of 100 bp
 * 10237 10967: contig of 731 bp in length
 * 10968 11067: gap of 100 bp
 * 11068 11787: contig of 720 bp in length
 * 11788 11887: gap of 100 bp
 * 11888 12655: contig of 768 bp in length
 * 12656 12755: gap of 100 bp
 * 12756 13497: contig of 742 bp in length
 * 13498 13597: gap of 100 bp
 * 13598 14365: contig of 768 bp in length
 * 14366 14465: gap of 100 bp
 * 14466 15203: contig of 738 bp in length
 * 15204 15303: gap of 100 bp
 * 15304 16010: contig of 707 bp in length
 * 16011 16110: gap of 100 bp
 * 16111 16872: contig of 762 bp in length
 * 16873 16972: gap of 100 bp
 * 16973 17715: contig of 743 bp in length
 * 17716 17815: gap of 100 bp
 * 17816 18587: contig of 772 bp in length
 * 18588 18687: gap of 100 bp
 * 18688 19464: contig of 777 bp in length
 * 19465 19564: gap of 100 bp
 * 19565 20303: contig of 739 bp in length
 * 20304 20403: gap of 100 bp
 * 20404 21135: contig of 732 bp in length
 * 21136 21235: gap of 100 bp

FEATURES
 source

Location/Qualifiers
 1. 50320
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="18"
 /map="18"

21236 21999: contig of 764 bp in length
 * 22000 22099: gap of 100 bp
 * 22100 22846: contig of 747 bp in length
 * 22847 22946: gap of 100 bp
 * 22947 23709: contig of 763 bp in length
 * 23710 23809: gap of 100 bp
 * 23810 24525: contig of 716 bp in length
 * 24526 24625: gap of 100 bp
 * 24626 25371: contig of 746 bp in length
 * 25372 25471: gap of 100 bp
 * 25472 26239: contig of 768 bp in length
 * 26240 26339: gap of 100 bp
 * 26340 27090: contig of 751 bp in length
 * 27091 27190: gap of 100 bp
 * 27191 27946: contig of 756 bp in length
 * 27947 28046: gap of 100 bp
 * 28047 28808: contig of 762 bp in length
 * 28809 28908: gap of 100 bp
 * 28909 29669: contig of 761 bp in length
 * 29670 29769: gap of 100 bp
 * 29770 30550: contig of 781 bp in length
 * 30551 30650: gap of 100 bp
 * 30651 31423: contig of 773 bp in length
 * 31424 31523: gap of 100 bp
 * 31524 32279: contig of 756 bp in length
 * 32280 32379: gap of 100 bp
 * 32380 33129: contig of 750 bp in length
 * 33130 33229: gap of 100 bp
 * 33230 33984: contig of 755 bp in length
 * 33985 34084: gap of 100 bp
 * 34085 34828: contig of 744 bp in length
 * 34829 34928: gap of 100 bp
 * 34929 35665: contig of 737 bp in length
 * 35666 35765: gap of 100 bp
 * 35766 36535: contig of 770 bp in length
 * 36536 36635: gap of 100 bp
 * 36636 37412: contig of 777 bp in length
 * 37413 37512: gap of 100 bp
 * 37513 38284: contig of 772 bp in length
 * 38285 38384: gap of 100 bp
 * 38385 39173: contig of 789 bp in length
 * 39174 39273: gap of 100 bp
 * 39274 40051: contig of 778 bp in length
 * 40052 40151: gap of 100 bp
 * 40152 40900: contig of 749 bp in length
 * 40901 41000: gap of 100 bp
 * 41001 41778: contig of 778 bp in length
 * 41779 41878: gap of 100 bp
 * 41879 42645: contig of 767 bp in length
 * 42646 42745: gap of 100 bp
 * 42746 43454: contig of 709 bp in length
 * 43455 43554: gap of 100 bp
 * 43555 44314: contig of 760 bp in length
 * 44315 44414: gap of 100 bp
 * 44415 45158: contig of 744 bp in length
 * 45159 45258: gap of 100 bp
 * 45259 46040: contig of 782 bp in length
 * 46041 46140: gap of 100 bp
 * 46141 46894: contig of 754 bp in length
 * 46895 46994: gap of 100 bp
 * 46995 47754: contig of 760 bp in length
 * 47755 47854: gap of 100 bp
 * 47855 48596: contig of 742 bp in length
 * 48597 48696: gap of 100 bp
 * 48697 49446: contig of 750 bp in length
 * 49447 49546: gap of 100 bp
 * 49547 50320: contig of 774 bp in length.

```

/clone="CTD-2307113"
/clone lib="CITD1 Human BAC"
BASE COUNT 12004 a 9574 c 10122 g 12507 t 6113 others
ORIGIN
Query Match 100.0%; Score 22; DB 2; Length 50320;
Best Local Similarity 100.0%; Pred. No. 7.7;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACGTTGACACCGCGGTTTCT 22
|||||
Db 567 GACGTTGACACCGCGGTTTCT 546

RESULT 7
AC016961
LOCUS
DEFINITION Homo sapiens chromosome 3 clone RP11-394J21, WORKING DRAFT
ACCESSION AC016961
VERSION AC016961.27 GI:21908412
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP; HTGS_ACTIVEFIN.
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 200947)
Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,
Albrooks,S.L., Amaratunge,H.C., Are,J.R., Ayele,M., Banks,T.,
Barbaria,J., Benton,J., Bimage,K., Blankenburg,K., Bonnin,D.,
Bouck,J., Bowie,S., Brieva,M., Brown,S., Brown,M., Bryant,N.P.,
Buhay,C., Burch,P., Burket,C., Burrell,K.L., Byrd,N.C.,
Carroll,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,
Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C.,
Cleveland,C.D., Davis,C., Coyle,M.D., Dathorne,S.R., David,R.,
Davila,M.L., Cox,C., Doy-Carroll,L., Dederich,D.A.,
Dellaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H.,
Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J.,
Earnhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M.,
Falls,I., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P.,
Gabis,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R.,
Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K.,
Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A., Hernandez,J.,
Hernandez,O., Hodgson,A., Hogue,M., Holloway,C., Hollins,B.,
Homsai,F., Howard,S., Huber,J., Hulyk,S., Hume,J., Jackson,L.E.,
Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S.,
Karissom,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C.,
Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L.,
Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W., Loubegeed,H.,
Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J.,
Maheshwari,M., Mapua,P., Martin,R., Martindale,A., Martinez,E.,
Massey,E., Mawhney,E., McLeod,M.P., Meador,M., Mei,G., Metzker,M.,
Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Morgan,M., Morris,S.,
Moser,M., Neal,D., Newton,J., Newton,N., Nguyen,A., Nguyen,N.,
Nguyen,N., Nickerson,E., Nwokkwo,S., Ogih,M., Okuwonu,G.,
Oragunye,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L.,
Peters,L., Pichens,A., Primus,E., Pu,L.L., Quiles,M., Ren,Y.,
Rives,M., Rojas,A., Rojuben,A., Ruiz,L., Ruiz,S., Savary,G.,
Scherer,S., Scott,G., Shen,H., Shoostari,N., Sisson,I.,
Sodergren,E., Sonaike,T., Sparks,A., Stanley,H., Stone,H.,
Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K., Tang,H.,
Tansley,J., Taylor,C., Taylor,T., Telford,B., Thomas,N., Thomas,S.,
Usmani,K., Vazquez,L., Vera,V., Villalon,D., Vinson,R., Wang,Q.,
Wang,S., Ward-Moore,S., Warren,R., Washington,C., Watlington,S.,
Williams,G., Williamson,A., Wleczyk,R., Wooden,S., Worley,K.,
Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,
Weinstock,G. and Gibbs,R.
Direct Submission
Unpublished
2 (bases 1 to 200947)
Worley,K.C.
Direct Submission
Submitted (09-DEC-1999) Human Genome Sequencing Center, Department

```

```

of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 200947)
Worley,K.C.
Direct Submission
Submitted (24-JUL-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Jul 19, 2002 this sequence version replaced gi:21539678.
----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Drafting Center Code: BCM
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: HNVB
Center clone name: RP11-394J21
----- Summary Statistics
Sequencing vector: M13;
Chemistry: Dye-terminator Big Dye; 98% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 200560 bases at least Q40
Consensus quality: 200582 bases at least Q30
Estimated insert size: 200621 bases at least Q20
Quality coverage: 8x in Q20 bases; sum-of-contigs estimation
-----
* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html)
* NOTE: This is a 'working draft' sequence. It currently
* consists of 2 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 13815: contig of 13815 bp in length
* 13816 13915: gap of unknown length
* 13916 200947: contig of 187032 bp in length.
Location/Qualifiers
1. 200947
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="3"
/clone="RP11-394J21"
BASE COUNT 59914 a 43317 c 42497 g 55118 t 101 others
ORIGIN
Query Match 100.0%; Score 22; DB 2; Length 200947;
Best Local Similarity 100.0%; Pred. No. 6.9;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GACGTTGACACCGCGGTTTCT 22
|||||
Db 75177 GACGTTGACACCGCGGTTTCT 75198

```

```

RESULT 8
AC094924/c
LOCUS
DEFINITION Rattus norvegicus clone CH230-3M13, *** SEQUENCING IN PROGRESS ***
ACCESSION AC094924
VERSION AC094924.3 GI:21716147
KEYWORDS HTG; HTGS_PHASE1.
SOURCE Norway rat.
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.

```

(see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).

* NOTE: This is a 'working draft' sequence. It currently consists of 48 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. * This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

1 1024: contig of 1024 bp in length
1025 1124: gap of unknown length
1125 2713: contig of 1589 bp in length
2714 2813: gap of unknown length
2814 4159: contig of 1346 bp in length
4160 4259: gap of unknown length
4260 5350: contig of 1091 bp in length
5351 5450: gap of unknown length
5451 6467: contig of 1017 bp in length
6468 6567: gap of unknown length
6568 7705: contig of 1138 bp in length
7706 7805: gap of unknown length
7806 8835: contig of 1030 bp in length
8836 8935: contig of unknown length
8936 10286: contig of 1351 bp in length
10287 10386: gap of unknown length
10387 11667: contig of 1281 bp in length
11668 11767: gap of unknown length
11768 13491: contig of 1724 bp in length
13492 13591: gap of unknown length
13592 14980: contig of 1389 bp in length
14981 15080: gap of unknown length
15081 16363: contig of 1283 bp in length
16364 16463: gap of unknown length
16464 18175: contig of 1712 bp in length
18176 18275: gap of unknown length
18276 19880: contig of 1605 bp in length
19881 19980: gap of unknown length
19981 21031: contig of 1051 bp in length
21032 21131: gap of unknown length
21132 22509: contig of 1378 bp in length
22510 22609: gap of unknown length
22610 24990: contig of 2381 bp in length
24991 25090: gap of unknown length
25091 26426: contig of 1336 bp in length
26427 26526: gap of unknown length
26527 28449: contig of 1923 bp in length
28450 28549: gap of unknown length
28550 30854: contig of 2305 bp in length
30855 30954: gap of unknown length
30955 32874: contig of 1920 bp in length
32875 32974: gap of unknown length
32975 35321: contig of 2347 bp in length
35322 35421: gap of unknown length
35422 36868: contig of 1447 bp in length
36869 36968: gap of unknown length
36969 39727: contig of 2759 bp in length
39728 39827: gap of unknown length
39828 40925: contig of 1098 bp in length
40926 41025: gap of unknown length
41026 44017: contig of 2992 bp in length
44018 44117: gap of unknown length
44118 45537: contig of 1420 bp in length
45538 45637: gap of unknown length
45638 46842: contig of 1105 bp in length
46843 48312: gap of unknown length
48313 48412: contig of 1470 bp in length
48413 50676: contig of 2264 bp in length
50677 50776: gap of unknown length
50778 52607: contig of 1831 bp in length
52608 52707: gap of unknown length
52708 54385: contig of 1678 bp in length
54386 54485: gap of unknown length

REFERENCE
AUTHORS
Muzay, D.M., Adams, C., Adio-Oduola, B., Ali-osman, F.R., Allen, C., Alsbrooks, S.L., Amarantunge, H.C., Are, J.R., Ayale, M., Banks, T., Barbaria, J., Benton, J., Binage, K., Blankenburg, K., Bonnin, D., Bouck, J., Bowie, S., Briviera, M., Brown, E., Brown, M., Bryant, N.P., Buha, C., Burch, P., Burkett, C., Burrell, K.L., Byrd, N.C., Carton, T.F., Carter, M., Cavazos, S.R., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Z., Chowdhry, I., Christopoulos, C., Cleveland, C.D., Cox, C., Coyle, M.D., Dathorne, S.R., David, R., Davila, M.L., Davis, C., Davy-Carroll, L., Dederich, D.A., Delaney, K.R., Delgado, O., Denn, A.L., Ding, Y., Dinh, H.H., DeLunay, K.J., Draper, H., Dugan-Rocha, S., Durbin, K.J., Eamhart, C., Edgar, D., Edwards, C.C., Elhaj, C., Escotto, M., Farns, T., Ferraguto, D., Flagg, N., Ford, J., Foster, P., Frantz, P., Gabisi, A., Gao, J., Garcia, A., Garner, T., Garza, N., Gill, R., Gorrell, J.H., Guevara, W., Gunaratne, P., Hale, S., Hamilton, K., Harris, C., Harris, K., Hart, M., Havlak, P., Hawes, A., Hernandez, J., Hernandez, O., Hodgson, A., Hogues, M., Holloway, C., Hollins, B., Homsi, F., Howard, S., Huber, J., Hulyk, S., Hume, J., Jackson, L.E., Jacobson, B., Jia, Y., Johnson, R., Jolivet, S., Joudah, S., Kelly, S., Khan, U., King, L., Korvah, J., Kovar, C., Kratovic, J., Kureshi, A., Landry, N., Leal, B., Lewis, L.C., Lewis, L., Li, J., Li, Z., Lichtarge, O., Lien, C., Liu, J., Liu, W., Loulsegged, H., Lozano, R.J., Lu, X., Lucier, A., Lucier, R., Luna, R., Ma, J., Maheshwari, M., Mapua, P., Martin, R., Martindale, A., Martinez, E., Massey, E., Mawhney, E., McLeod, M.P., Meador, M., Mei, G., Metzker, M., Miner, G., Miner, Z., Mitchell, T., Mohabbat, K., Morgan, M., Morris, S., Moser, M., Neal, D., Newton, J., Newton, N., Nguyen, A., Nguyen, N., Nguyen, N., Nickerson, E., Nwokenwo, S., Ogih, M., Okwuonu, G., Oragunye, N., Oviedo, R., Pace, A., Payton, B., Peery, J., Perez, L., Peters, L., Pickens, R., Primus, E., Pu, L.L., Quiles, M., Ren, Y., Rives, M., Rojas, A., Rojubokan, I., Rolfe, M., Ruiz, S., Savary, G., Scherer, S., Scott, G., Shen, H., Shoohtari, N., Sisson, I., Sodergren, E., Sonaite, T., Sparks, A., Stanley, H., Stone, H., Sutton, A., Svatek, A., Tabor, P., Tamerisa, A., Tamerisa, K., Tang, H., Tansey, J., Taylor, C., Taylor, T., Telford, B., Thomas, N., Thomas, S., Usmani, K., Vasquez, L., Vera, V., Villalon, D., Vinson, R., Wang, Q., Wang, S., Ward-Moore, S., Warren, R., Washington, C., Watlington, S., Williams, G., Williamson, A., Wlezyk, R., Wooden, S., Worley, K., Wu, C., Wu, Y., Wu, Y.F., Zhou, J., Zorrilla, S., Nelson, D., Weinstein, G., and Gibbs, R.

2 (bases 1 to 124849)
Unpublished
Worley, K.C.
Direct Submission
Submitted (15-SEP-2001) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 124849)
Worley, K.C.
Direct Submission
Submitted (10-JUL-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
On Jul 9, 2002 this sequence version replaced gi:17941723.
----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GBSB
Center Clone name: CH230-3M13
----- Summary Statistics
Sequencing vector: Plasmid
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 78224 bases at least Q40
Consensus quality: 82998 bases at least Q30
Consensus quality: 86973 bases at least Q20

* NOTE: Estimated insert size may differ from sequence length

```

* 54486 58259: contig of 3774 bp in length
* 58260 61238: gap of unknown length
* 58360 61238: contig of 2869 bp in length
* 61329 64029: gap of unknown length
* 61329 64029: contig of 2701 bp in length
* 64030 64129: gap of unknown length
* 64130 67245: contig of 3116 bp in length
* 67245 70188: gap of unknown length
* 70188 70289: contig of 2843 bp in length
* 70289 73911: contig of 3623 bp in length
* 73911 74011: gap of unknown length
* 74011 78470: contig of 4459 bp in length
* 78470 78570: gap of unknown length
* 78570 82831: contig of 4261 bp in length
* 82831 82932: gap of unknown length
* 82932 86995: contig of 3964 bp in length
* 86995 91366: gap of unknown length
* 91366 91467: contig of 4371 bp in length
* 91467 95236: gap of unknown length
* 95236 95337: contig of 3770 bp in length
* 95337 98740: gap of unknown length
* 98740 102540: contig of 3404 bp in length
* 102540 102641: gap of unknown length
* 102641 109458: contig of 3700 bp in length
* 109458 109557: gap of unknown length
* 109558 117662: contig of 6817 bp in length
* 117662 117663: contig of 8105 bp in length
* 117663 124849: gap of unknown length
* 124849 124849: contig of 7087 bp in length.

```

FEATURES

Location/Qualifiers

```

1..124849
/organism="Rattus norvegicus"
/db_xref="taxon:10116"
/c clones="CH230-3M13"
BASE COUNT 31267 a 26760 c 26699 g 30221 t 9902 others
ORIGIN

```

```

Query Match 92.7%; Score 20.4; DB 2; Length 124849;
Best Local Similarity 95.5%; Pred. No. 41;
Matches 21; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GACGTTGACACGGCGGTTTCT 22
Db 36120 GACGTTGACACGGCGGTTTCT 36099

```

RESULT 9

```

AC126477
LOCUS Rattus norvegicus clone CH230-43M20, linear HTG 11-JUL-2002
DEFINITION *** 62 unordered pieces.

```

AC126477 161691 bp DNA

AC126477.1 GI:21700419

HTG; HTGS PHASE1.

Norway rat.

Rattus norvegicus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

1 (bases 1 to 161691)

Muzny, D.M., Adams, C., Adio-Oduola, B., Ali-oesan, F.R., Allen, C., Alsbrooks, S.L., Amaratunge, H.C., Are, J.R., Ayelle, M., Banks, T., Barbarta, J., Benton, J., Binage, K., Blankenburg, K., Bonnin, D., Bouck, J., Bowie, S., Brieva, M., Brown, E., Brown, M., Bryant, N.P., Buhay, C., Burch, P., Burkett, C., Burrell, K.L., Byrd, N.C., Carron, F.F., Carter, M., Cavazos, S.R., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Z., Chowdhry, I., Christopoulos, C., Cleveland, C.D., Cox, C., Coyle, M.D., Dathorne, S.R., David, R., Davila, M.L., Davis, C., Davy-Carroll, L., Dederich, D.A., Delaney, K.R., Delgado, O., Denna, A.L., Ding, Y., Dinh, H.H., Douthwaite, K.J., Draper, H., Dugan-Rocha, S., Durbin, K.J.,

Barnhart, C., Edgar, D., Edwards, C.C., Elhaj, C., Escotto, M., Falls, T., Ferraguto, D., Flagg, N., Ford, J., Foster, P., Frantz, P., Gabisi, A., Gao, J., Garcia, A., Garner, T., Garza, N., Gill, R., Gorrill, J.H., Guevara, W., Gunaratne, P., Hale, S., Hamilton, K., Harris, C., Harris, K., Hart, M., Havlak, P., Hawes, A., Hernandez, J., Hernandez, O., Hodgson, A., Hogue, M., Holloway, C., Hollings, B., Homs, F., Howard, S., Huber, J., Huiyk, S., Hume, J., Jackson, L.E., Jacobson, B., Jia, Y., Johnson, R., Jolivet, S., Joudah, S., Karlsson, E., Kelly, S., Khan, U., King, L., Korvah, J., Kovar, C., Kratovic, J., Kureshi, A., Landry, N., Leal, B., Lewis, L.C., Lewis, L., Li, J., Li, Z., Lichtarge, O., Lieu, C., Liu, J., Liu, W., Louie, H., Lozano, R.J., Lu, X., Lucier, A., Lucier, R., Luna, R., Ma, J., Maheshwari, M., Mapua, P., Martin, R., Martindale, A., Martinez, E., Massey, E., Mawhney, E., McLeod, M.P., Meador, M., Mei, G., Metzger, M., Miner, G., Miner, Z., Mitchell, T., Mohabbat, K., Morgan, M., Morris, S., Moser, M., Neal, D., Newton, J., Newton, N., Nguyen, A., Nguyen, N., Nguyen, N., Nickerson, E., Nwokenkwo, S., Ogih, M., Okwuonu, G., Oragunye, N., Oviedo, R., Pace, A., Payton, B., Peery, J., Perez, L., Peters, L., Pickens, R., Primus, E., Pu, L.L., Quiles, M., Ren, Y., Rives, M., Rojas, A., Rojubokan, I., Rolfe, M., Ruiz, S., Savary, G., Scherer, S., Scott, G., Shen, H., Shoohtari, N., Sisson, I., Sodergren, E., Sonaik, T., Sparks, A., Stanley, H., Stone, H., Sutton, A., Svatek, A., Tabors, P., Tamerisa, A., Tamerisa, K., Tang, H., Tansey, J., Taylor, C., Taylor, T., Telford, B., Thomas, N., Thomas, S., Usmani, K., Vasquez, L., Vera, V., Villalon, D., Vinson, R., Wang, Q., Wang, S., Ward-Moore, S., Warren, R., Washington, C., Watlington, S., Williams, G., Williamson, A., Wleczyk, R., Wooden, S., Worley, K., Wu, C., Wu, Y., Wu, Y.F., Zhou, J., Zorrilla, S., Nelson, D., Weinstein, G., and Gibbs, R.

Direct Submission
Unpublished
2 (bases 1 to 161691)
Worley, K.C.
Direct Submission
Submitted (06-JUL-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 161691)
Worley, K.C.
Direct Submission
Submitted (11-JUL-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu

Project Information
Center project name: GFBV
Center clone name: CH230-43M20

Summary Statistics
Sequencing vector: Plasmid;
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 112006 bases at least Q40
Consensus quality: 118246 bases at least Q30
Consensus quality: 122772 bases at least Q20

* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 62 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 1016: contig of 1016 bp in length
* 1017 1116: gap of unknown length
* 1117 2386: contig of 1170 bp in length
* 2287 2386: gap of unknown length

73444: gap of unknown length
73445: contig of 3147 bp in length
73446: gap of unknown length
73447: contig of 2835 bp in length
73448: gap of unknown length
73449: contig of 3228 bp in length
73450: gap of unknown length
73451: contig of 3429 bp in length
73452: gap of unknown length
73453: contig of 2228 bp in length
73454: gap of unknown length
73455: contig of 2173 bp in length
73456: gap of unknown length
73457: contig of 2802 bp in length
73458: gap of unknown length
73459: contig of 3511 bp in length
73460: gap of unknown length
73461: contig of 2858 bp in length
73462: gap of unknown length
73463: contig of 3594 bp in length
73464: gap of unknown length
73465: contig of 3866 bp in length
73466: gap of unknown length
73467: contig of 5078 bp in length
73468: gap of unknown length
73469: contig of 2469 bp in length
73470: gap of unknown length
73471: contig of 2463 bp in length
73472: gap of unknown length
73473: contig of 4073 bp in length
73474: gap of unknown length
73475: contig of 161691 bp in length

Query Match 92.7%; Score 20.4; DB 2; Length 161691;
Best Local Similarity 95.5%; Pred. No. 41;
Matches 21; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GACGTGACACGCGGTTCCT 22
DB 148271 GACGTGACACGCGGTTCCT 148292

RESULT 10
AL596177/c
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

AL596177
Human DNA sequence from clone RP11-325P15 on chromosome 1, complete sequence.
AL596177.4 GI:15552942
HTG.
human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 91084)
Direct Submission
Submitted (08-SEP-2001) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk
requests: clonerequest@sanger.ac.uk
On Sep 11, 2001 this sequence version replaced gi:15022347.
During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.
This sequence was finished as follows unless otherwise noted: all
regions were either double-stranded or sequenced with an alternate
chemistry or covered by high quality data (i.e. phred quality >= 30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by at least
one plasmid subclone or more than one M13 subclone; and the
assembly was confirmed by restriction digest. The following

4099: contig of 1713 bp in length
4199: gap of unknown length
4200: contig of 1529 bp in length
5828: gap of unknown length
7407: contig of 1579 bp in length
7507: gap of unknown length
7509: contig of 1252 bp in length
8759: gap of unknown length
8760: contig of 1547 bp in length
8860: gap of unknown length
10407: contig of 1464 bp in length
11970: gap of unknown length
12070: contig of 1219 bp in length
13289: gap of unknown length
13389: contig of 1088 bp in length
14477: gap of unknown length
14577: contig of 1386 bp in length
15963: gap of unknown length
16063: contig of 1394 bp in length
17457: gap of unknown length
17557: contig of 1958 bp in length
19515: gap of unknown length
19615: contig of 1370 bp in length
20985: gap of unknown length
21085: contig of 1955 bp in length
23040: gap of unknown length
23140: contig of 1932 bp in length
25072: gap of unknown length
25172: contig of 2858 bp in length
28030: gap of unknown length
28130: contig of 1985 bp in length
30113: gap of unknown length
30215: contig of 1433 bp in length
31648: gap of unknown length
31748: contig of 1640 bp in length
33488: gap of unknown length
33522: contig of 1764 bp in length
33523: gap of unknown length
36783: contig of 1431 bp in length
36883: gap of unknown length
37977: contig of 1094 bp in length
38077: gap of unknown length
40083: contig of 2006 bp in length
40183: gap of unknown length
41999: contig of 1816 bp in length
42099: gap of unknown length
44664: contig of 2565 bp in length
44764: gap of unknown length
45960: contig of 1196 bp in length
46060: gap of unknown length
47229: contig of 1669 bp in length
47829: gap of unknown length
49786: contig of 1957 bp in length
49886: gap of unknown length
51125: contig of 1239 bp in length
51225: gap of unknown length
54559: contig of 3234 bp in length
54559: gap of unknown length
56015: contig of 1456 bp in length
56115: gap of unknown length
59017: contig of 2902 bp in length
59117: gap of unknown length
61064: contig of 1947 bp in length
61164: gap of unknown length
63433: contig of 2269 bp in length
63533: gap of unknown length
65766: contig of 2233 bp in length
65866: gap of unknown length
68115: contig of 2249 bp in length
68215: gap of unknown length
70969: contig of 2754 bp in length
71069: gap of unknown length
73344: contig of 2275 bp in length

abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em., EMBL; Sw., SWISSPROT; Tr., TREMBL; Wp., WORMPEP; information on the WORMPEP database can be found at http://www.sanger.ac.uk/Projects/C_elegans/wormpep. This sequence was generated from part of bacterial clone contigs of human chromosome 1, constructed by the Sanger Centre Chromosome 1 Mapping Group. Further information can be found at <http://www.sanger.ac.uk/HGP/Chrl>. RP11-325P15 is from the library RCI-11.2 constructed by the group of Pieter de Jong. For further details see <http://www.chori.org/bacpac/home.htm>.
VECTOR: pBAC3.6

IMPORTANT: This sequence is not the entire insert of clone RP11-325P15. It may be shorter because we sequence overlapping sections only once, except for a short overlap. The true left end of clone RP11-325P15 is at 1 in this sequence. The true left end of clone RP11-337C18 is at 89085 in this sequence. The true right end of clone RP4-704D21 is at 41034 in this sequence.

```

FEATURES             Location/Qualifiers
     source            1..91084
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="11"
                     /clone="RP11-325P15"
                     /clone_lib="RCI-11.2"
BASE COUNT          27074 a 19586 c 19377 g 25047 t
ORIGIN

```

```

Query Match          85.5%; Score 18.8; DB 9; Length 91084;
Best Local Similarity 90.9%; Pred. No. 2.5e+02;
Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

```

Qy 1 GACGTGACACGGCGGTTC 22
Db 9827 GACGTGACACACGAGTGTTC 9806

```

```

RESULT 11
LOCUS              AC026579
DEFINITION          Homo sapiens chromosome 15 clone RP11-76D16 map 15, LOW-PASS
                     SEQUENCE SAMPLING.
ACCESSION            AC026579
VERSION              AC026579.2 GI:17048293
KEYWORDS              HTG; HTGS PHASE0.
SOURCE               Homo sapiens
ORGANISM             Homo sapiens
                     Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
                     Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE            1 (bases 1 to 112259)
AUTHORS              Birren, B., Linton, L., Nusbaum, C., Lander, E., Abramson, H., Allen, N.,
                     Anderson, S., Baldwin, J., Barna, N., Bastien, V., Bedalov, P.,
                     Boguslavsky, L., Bouckge, B., Brown, A., Burkett, G.,
                     Campolongo, A., Castelle, P., Chao, P., Colangelo, M., Collins, S.,
                     Collymore, A., Cooke, P., DeArnell, K., Dewar, K., Diaz, J.S.,
                     Dodge, S., Domino, M., Doyle, M., Ferreira, P., FitzHugh, W., Gage, D.,
                     Galagan, J., Gardyna, S., Ginde, S., Goyette, M., Graham, L.,
                     Grand-Pierre, N., Grant, C., Hagos, B., Heaford, A., Horton, L.,
                     Howland, J.C., Iliev, I., Johnson, R., Jones, C., Kann, L., Karatas, A.,
                     Klein, J., LaRocque, K., Lamazares, R., Landers, T., Lehoczy, J.,
                     Levine, R., Lieu, C., Liu, G., Locke, K., Macdonald, P., Marquis, N.,
                     McCarthy, M., McEwan, P., McGurk, A., McKernan, K., McPheeters, R.,
                     Melarini, J., Meneus, L., Mihova, T., Miranda, C., Mlenga, V., Morrow, J.,
                     Murphy, T., Naylor, J., Norman, C.H., O'Connor, T., O'Donnell, P.,
                     O'Neill, D., Olivarez, T.M., Oliver, J., Peterson, K., Pierre, N.,
                     Pisani, C., Pollara, V., Raymond, C., Riley, R., Rogov, P., Rothman, D.,
                     Roy, A., Santos, R., Schauer, S., Severy, P., Spencer, B.,
                     Birren, B., Linton, L., Nusbaum, C., Lander, E., Abramson, H., Allen, N.,
                     Anderson, S., Baldwin, J., Barna, N., Bastien, V., Bedalov, P.,
                     Boguslavsky, L., Bouckge, B., Brown, A., Burkett, G.,
                     Campolongo, A., Castelle, P., Chao, P., Colangelo, M., Collins, S.,
                     Collymore, A., Cooke, P., DeArnell, K., Dewar, K., Diaz, J.S.,
                     Dodge, S., Domino, M., Doyle, M., Ferreira, P., FitzHugh, W., Gage, D.,
                     Galagan, J., Gardyna, S., Ginde, S., Goyette, M., Graham, L.,
                     Grand-Pierre, N., Grant, C., Hagos, B., Heaford, A., Horton, L.,
                     Howland, J.C., Iliev, I., Johnson, R., Jones, C., Kann, L., Karatas, A.,
                     Klein, J., LaRocque, K., Lamazares, R., Landers, T., Lehoczy, J.,
                     Levine, R., Lieu, C., Liu, G., Locke, K., Macdonald, P., Marquis, N.,
                     McCarthy, M., McEwan, P., McGurk, A., McKernan, K., McPheeters, R.,
                     Melarini, J., Meneus, L., Mihova, T., Miranda, C., Mlenga, V., Morrow, J.,
                     Murphy, T., Naylor, J., Norman, C.H., O'Connor, T., O'Donnell, P.,
                     O'Neill, D., Olivarez, T.M., Oliver, J., Peterson, K., Pierre, N.,
                     Pisani, C., Pollara, V., Raymond, C., Riley, R., Rogov, P., Rothman, D.,
                     Roy, A., Santos, R., Schauer, S., Severy, P., Spencer, B.,

```

Stange-Thomann, N., Stojanovic, N., Subramanian, A., Talamas, J.,
Teefaye, S., Theodore, J., Tirrell, A., Travers, M., Trigilio, J.,
Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J.,
Young, G., Zainoun, J., Zimmer, A. and Zody, M.
Direct Submission
Submitted (22-MAR-2000) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Nov 22, 2001 this sequence version replaced gi:7284066.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: <http://www-seq.wi.mit.edu>
Contact: sequence.submissions@genome.wi.mit.edu
----- Project Information
Center project name: L8359
Center clone name: 76_D_16

* NOTE: This record contains 139 individual
* sequencing reads that have not been assembled into
* contigs. Runs of N are used to separate the reads
* and the order in which they appear is completely
* arbitrary. Low-pass sequence sampling is useful for
* identifying clones that may be gene-rich and allows
* overlap relationships among clones to be deduced.
* However, it should not be assumed that this clone
* will be sequenced to completion. In the event that
* the record is updated, the accession number will
* be preserved.

```

1 715 814: gap of 100 bp
815 1541: contig of 727 bp in length
1542 1641: gap of 100 bp
1642 2329: contig of 688 bp in length
2330 2429: gap of 100 bp
2430 3143: contig of 714 bp in length
3144 3243: gap of 100 bp
3244 3955: contig of 712 bp in length
3956 4055: gap of 100 bp
4056 4726: contig of 671 bp in length
4727 4826: gap of 100 bp
4827 5533: contig of 707 bp in length
5534 5633: gap of 100 bp
5634 6329: contig of 696 bp in length
6330 6429: gap of 100 bp
6430 7152: contig of 723 bp in length
7153 7252: gap of 100 bp
7253 7978: contig of 726 bp in length
7979 8078: gap of 100 bp
8079 8772: contig of 694 bp in length
8773 8872: gap of 100 bp
8873 9570: contig of 698 bp in length
9571 9670: gap of 100 bp
9671 10355: contig of 685 bp in length
10356 10455: gap of 100 bp
10456 11162: contig of 707 bp in length
11163 11262: gap of 100 bp
11263 11862: contig of 700 bp in length
11863 12062: gap of 100 bp
12063 12759: contig of 697 bp in length
12760 12859: gap of 100 bp
12860 13633: contig of 674 bp in length
13634 13633: gap of 100 bp
13634 14355: contig of 722 bp in length
14356 14455: gap of 100 bp
14456 15160: contig of 705 bp in length
15161 15260: gap of 100 bp
15261 15972: contig of 712 bp in length
15973 16072: gap of 100 bp
16073 16778: contig of 706 bp in length
16779 16878: gap of 100 bp

```

* 16879 17563: contig of 685 bp in length
 * 17564 17663: gap of 100 bp
 * 17664 18372: contig of 709 bp in length
 * 18373 18472: gap of 100 bp
 * 18473 19195: contig of 723 bp in length
 * 19196 19295: gap of 100 bp
 * 19296 19984: contig of 689 bp in length
 * 19985 20084: gap of 100 bp
 * 20085 20778: contig of 694 bp in length
 * 20779 20878: gap of 100 bp
 * 20879 21570: contig of 692 bp in length
 * 21571 21670: gap of 100 bp
 * 21671 22319: contig of 649 bp in length
 * 22320 22419: gap of 100 bp
 * 22420 23122: contig of 703 bp in length
 * 23123 23222: gap of 100 bp
 * 23223 23937: contig of 715 bp in length
 * 23938 24037: gap of 100 bp
 * 24038 24760: contig of 723 bp in length
 * 24761 24860: gap of 100 bp
 * 24861 25564: contig of 704 bp in length
 * 25565 25664: gap of 100 bp
 * 25665 26372: contig of 708 bp in length
 * 26373 26472: gap of 100 bp
 * 26473 27155: contig of 683 bp in length
 * 27156 27255: gap of 100 bp
 * 27256 27945: contig of 690 bp in length
 * 27946 28045: gap of 100 bp
 * 28046 28753: contig of 708 bp in length
 * 28754 28853: gap of 100 bp
 * 28854 29563: contig of 710 bp in length
 * 29564 29663: gap of 100 bp
 * 29664 30381: contig of 718 bp in length
 * 30382 30481: gap of 100 bp
 * 30483 31201: contig of 720 bp in length
 * 31202 31301: gap of 100 bp
 * 31302 31988: contig of 687 bp in length
 * 31989 32088: gap of 100 bp
 * 32089 32777: contig of 689 bp in length
 * 32778 32877: gap of 100 bp
 * 32878 33574: contig of 697 bp in length
 * 33575 33674: gap of 100 bp
 * 33675 34379: contig of 705 bp in length
 * 34380 34479: gap of 100 bp
 * 34480 35191: contig of 712 bp in length
 * 35192 35291: gap of 100 bp
 * 35292 36001: contig of 710 bp in length
 * 36002 36101: gap of 100 bp
 * 36102 36806: contig of 705 bp in length
 * 36807 36906: gap of 100 bp
 * 36907 37615: contig of 709 bp in length
 * 37616 37715: gap of 100 bp
 * 37716 38396: contig of 681 bp in length
 * 38397 38496: gap of 100 bp
 * 38497 39187: contig of 691 bp in length
 * 39188 39287: gap of 100 bp
 * 39288 39986: contig of 699 bp in length
 * 39987 40086: gap of 100 bp
 * 40087 40788: contig of 702 bp in length
 * 40789 40888: gap of 100 bp
 * 40889 41603: contig of 715 bp in length
 * 41604 41703: gap of 100 bp
 * 41704 42424: contig of 721 bp in length
 * 42425 42524: gap of 100 bp
 * 42525 43231: contig of 707 bp in length
 * 43232 43331: gap of 100 bp
 * 43332 44029: contig of 698 bp in length
 * 44030 44129: gap of 100 bp
 * 44130 44815: contig of 686 bp in length
 * 44816 44915: gap of 100 bp
 * 44916 45611: contig of 696 bp in length
 * 45612 45711: gap of 100 bp
 * 45712 46403: contig of 692 bp in length

* 46404 46503: gap of 100 bp
 * 46504 47206: contig of 703 bp in length
 * 47207 47306: gap of 100 bp
 * 47307 47995: contig of 689 bp in length
 * 47996 48095: gap of 100 bp
 * 48096 48793: contig of 698 bp in length
 * 48794 48893: gap of 100 bp
 * 48894 49463: contig of 570 bp in length
 * 49464 49563: gap of 100 bp
 * 49564 50250: contig of 687 bp in length
 * 50251 50350: gap of 100 bp
 * 50351 51091: contig of 741 bp in length
 * 51092 51191: gap of 100 bp
 * 51192 51909: contig of 718 bp in length
 * 51910 52009: gap of 100 bp
 * 52010 52713: contig of 704 bp in length
 * 52714 52813: gap of 100 bp
 * 52814 53526: contig of 713 bp in length
 * 53527 53626: gap of 100 bp
 * 53627 54337: contig of 711 bp in length
 * 54338 54437: gap of 100 bp
 * 54438 55155: contig of 718 bp in length

Query Match 85.5%; Score 18.8; DB 2; Length 112259;

Best Local Similarity 90.9%; Pred. No. 2.4e+02;

Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 GACGTTGACAAACGGCGGTTTCT 22
 Db 77471 GACGTTGACAAACGGCGTCTCT 77492

RESULT 12

AC020629/c

LOCUS Homo sapiens 12q BAC RP11-76E16 (Roswell Park Cancer Institute

DEFINITION Human BAC Library) complete sequence.

ACCESSION AC020629.6 GI:7656675

VERSION AC020629.6

KEYWORDS HTG.

SOURCE Homo sapiens.

ORGANISM Homo sapiens.

REFERENCE

AUTHORS

142971 bp DNA linear PRI 07-MAR-2002
 Homo sapiens 12q BAC RP11-76E16 (Roswell Park Cancer Institute
 Human BAC Library) complete sequence.

AC020629
 Homo sapiens 12q BAC RP11-76E16 (Roswell Park Cancer Institute
 Human BAC Library) complete sequence.
 AC020629.6 GI:7656675
 HTG.
 Homo sapiens.
 Homo sapiens.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 142971)
 Muzny, D.M., Adams, C., Bailey, M., Barbaria, J., Blankenburg, K.,
 Bodota, B., Bouch, J., Bowie, S., Brooks, A., Buhay, C., Bunac, C.,
 Burkett, C., Burrows, J., Carter, M., Chacko, J., Chen, Z., Cox, C.,
 David, R., Delgado, O., Deshazo, D., Ding, Y., Donah-Rashid, N.,
 Dugan-Rocha, S., Durbin, K.J., Fernandez, C., Ferraguto, D.,
 Forcum-Tansey, J., Frantz, P., Ganes, R., Gorrell, J.H., Gorrell, L.L.,
 Guevara, W., Harris, K., Hernandez, J., Hodgson, A., Hognes, M., Jia, Y.,
 Holloway, C., Hosak, H., Issar, A., Jackson, L.E., Jackson, L., Kovar, C.,
 Jones, M., Kelly, S., Kneitz, S., Kondejewski, N., Kong, Y., Liu, W.,
 Lau, S., Leal, B., Lee, E., Li, Z., Lucier, R., Marondel, I., Martin, R.,
 Logan, O., Lozano, R.J., Lu, J., Lucier, R., Marondel, I., Martin, R.,
 Martinez, C., McLeod, M.P., Mei, G., Merscher, S., Nash, S.,
 Montgomery, K.T., Morgan, M., Morris, S., Nash, S., Nash, S.,
 Nguyen, B., Nguyen, N., Nguyen, S., Osval, G., Parish, B., Paxton, S.,
 Payton, B., Perez, L., Pu, L.L., Quiles, M., Reiter, D., Rives, M.,
 Samuel, S., Say, J.J., Scherer, S., Shah, E., Shen, H., Shim, C., Simon, M.,
 Sparks, A., Stamps, A., Sugang, R., Tabor, P., Taylor, T., Vasquez, L.,
 Vinson, R., Vo, Q., Wahab, M., Watlington, S., Weinstein, G.,
 Weinstein, I.R., Williamson, A., Worley, K., Wren, J., Wrensford, G.,
 Xiang, A.M., Yang, R., Yu, W., Zhou, X., Kucheraipati, R., Nelson, D. and
 Gibbs, R.

Direct Submission

Unpublished

2 (bases 1 to 142971)

Worley, K.C.

Direct Submission

Submitted (07-JAN-2000) Human Genome Sequencing Center, Department

of Molecular and Human Genetics, Baylor College of Medicine, One

REFERENCE
AUTHORS
TITLE
JOURNAL

Baylor Plaza, Houston, TX 77030, USA
Worley, K.C.
Direct Submission
Submitted (27-APR-2000) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
4 (bases 1 to 142971)
Worley, K.C.
Direct Submission
Submitted (28-APR-2000) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
5 (bases 1 to 142971)
Worley, K.C.
Direct Submission
Submitted (07-MAR-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Apr 27, 2000 this sequence version replaced gi:7025656.
INFORMATION: <http://www.hgsc.bcm.tmc.edu/> or email
gc-help@bcm.tmc.edu

COMMENT

CLONE LENGTH: This sequence does not necessarily represent the
entire insert of this clone. Overlapping regions of clones are only
sequenced and submitted once, so the sequence for the remainder of
the insert may be found in the record for the adjacent clones.
Overlapping clones are noted at the beginning and end of the
Features listing.

ANNOTATION OF FEATURES:

STSs are identified using ePCR (Genome Res. 7:541-550) searches
of a local database that includes entries from dbSTS, GDB, and
local mapping efforts.

Repeats are identified using RepeatMasker (A. Smit and P. Green,
unpublished.) for Human and Mouse sequences.
Genes and Region of sequence similarity are identified by BLAST
(Nuc. Acids Res 25:3389-3402) similarity (expect < 1e-34) to the
EST and cDNA sequences. Genes demonstrate at least two exons
flanked by consensus splice sites that maintained sequence
continuity across the splice junctions. Sequences that are not
identical matches are annotated as similar.

SEQUENCING READ COVERAGE: Sequencing is completed to a minimum
standard of double strand coverage with a minimum of 2 clones and 2
reads with no ambiguities or 2 chemistries with a minimum of 2
clones and 3 reads with no ambiguities. If the sequence quality for
a region does not meet this standard, it will be indicated in the
annotation as Low Coverage.

QUALITY OF INDIVIDUAL BASES: This sequence meets stringent quality
standards - estimated error rate less than 1 per 10,000 bases.
Reports of lowest quality individual bases and measures of base
quality are listed below. Description of the metrics can be found
at URL: <http://gc.bcm.tmc.edu:8088/quality.info/genbank.annotation.html>.

QUALSTAT-REPORT-----

```

----- Summary Statistics -----
Contig length: 142971
Phrap values in estimate: 140680
Average error rate (BCM-Phrap estimate): 0.000118703
Fraction of Phrap values less than 40 : 0.0433608
Number of consensus changing edits: 31
Number of N's in consensus : 0

----- Consensus changing edits -----
Position Original+Context Edited+Context
11171 gctctctgga(n)caaaaaaaa gctctctgga(a)caaaaaaaa
12018 tatctatctg(n)ctatatctat tatctatctg(t)ctatatctat

```

Distribution of Quality < 40 Bases		-----										
#	bases	1000	900	800	700	600	500	400	300	200	100	0
16604	16604	1000	900	800	700	600	500	400	300	200	100	0
18385	18385	1000	900	800	700	600	500	400	300	200	100	0
21879	21879	1000	900	800	700	600	500	400	300	200	100	0
21880	21880	1000	900	800	700	600	500	400	300	200	100	0
21881	21881	1000	900	800	700	600	500	400	300	200	100	0
21906	21906	1000	900	800	700	600	500	400	300	200	100	0
22169	22169	1000	900	800	700	600	500	400	300	200	100	0
38915	38915	1000	900	800	700	600	500	400	300	200	100	0
55675	55675	1000	900	800	700	600	500	400	300	200	100	0
57658	57658	1000	900	800	700	600	500	400	300	200	100	0
74371	74371	1000	900	800	700	600	500	400	300	200	100	0
80221	80221	1000	900	800	700	600	500	400	300	200	100	0
88248	88248	1000	900	800	700	600	500	400	300	200	100	0
94121	94121	1000	900	800	700	600	500	400	300	200	100	0
95228	95228	1000	900	800	700	600	500	400	300	200	100	0
95230	95230	1000	900	800	700	600	500	400	300	200	100	0
95231	95231	1000	900	800	700	600	500	400	300	200	100	0
104931	104931	1000	900	800	700	600	500	400	300	200	100	0
118168	118168	1000	900	800	700	600	500	400	300	200	100	0
120974	120974	1000	900	800	700	600	500	400	300	200	100	0
120976	120976	1000	900	800	700	600	500	400	300	200	100	0
120978	120978	1000	900	800	700	600	500	400	300	200	100	0
120980	120980	1000	900	800	700	600	500	400	300	200	100	0
139505	139505	1000	900	800	700	600	500	400	300	200	100	0
139508	139508	1000	900	800	700	600	500	400	300	200	100	0
139598	139598	1000	900	800	700	600	500	400	300	200	100	0
140104	140104	1000	900	800	700	600	500	400	300	200	100	0
140149	140149	1000	900	800	700	600	500	400	300	200	100	0

Version: 1.01 qxfo.
Location/Qualifiers
1. 142971
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="12"
/clone="RP11-76E16"
147, 236
/rpt_family="A-rich"
258, 378
/rpt_family="AluJb"
complement(537, 840)
/rpt_family="AluSx"
complement(1722, 2137)
/rpt_family="Tigger2a"
complement(3161, 3584)
/rpt_family="MSTD"
3228, 3502
/standard_name="G05956"
/db_xref="dbSTS:11453"
3347, 3490
/standard_name="Cda0yb04"
/db_xref="GDB:446196"

/organism="Oryza sativa (japonica cultivar-group)"

/db xref="taxon:39947"

/chromosome="11"

/clone="OSUNBa0038P07"

BASE COUNT 49578 a 37891 c 37968 g 48745 t 441 others
ORIGIN

Query Match 85.5%; Score 18.8; DB 2; Length 174623;
Best Local Similarity 90.9%; Pred. No. 2.3e+02;
Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Oy 1 GACGTTGACACGGCGGTTTCT 22

Db 150055 GACGTTGACACGGCGGTTTCT 150076

RESULT 14

AC1015706/c

LOCUS

DEFINITION

SEQUENCE, 17 unordered pieces.

AC1015706

AC1015706.3 GI:8096833

HTG; HTGS_PHASE1; HTGS_DRAFT.

KEYWORDS

SOURCE

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

misc_feature 3585..3735

/note="Region similar to Hs#S936143 aJ27D02.s1 Homo

sapiens cDNA: AA789222"

repeat_region 3885..3910

/rpt_family="(TA)n"

repeat_region complement(3917..4030)

/rpt_family="AluJb"

repeat_region complement(4544..4681)

/rpt_family="MIR"

misc_feature 5145..5510

/note="Region similar to Hs#S953417 oa06g05.s1 Homo

sapiens cDNA: AA765439"

misc_feature 5553..5859

/note="Region similar to Hs#S1195593 qt69f10.x1 Homo

sapiens cDNA: AI285403"

Query Match 85.5%; Score 18.8; DB 9; Length 142971;

Best Local Similarity 90.9%; Pred. No. 2.4e+02;

Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Oy 1 GACGTTGACACGGCGGTTTCT 22

Db 133396 GACGTTGACACGGCGGTTTCT 133375

RESULT 13

AC108224

LOCUS

DEFINITION

Oryza sativa (japonica cultivar-group) chromosome 11 clone

OSUNBa0038P07, *** SEQUENCING IN PROGRESS ***, 4 ordered pieces.

AC108224

AC108224.3 GI:21392687

HTG; HTGS_PHASE2.

SOURCE

ORGANISM

Oryza sativa (japonica cultivar-group).

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;

Ehrhartoideae; Oryzoae; Oryza.

1 (bases 1 to 174623)

Sharma,T.R., McCombie,W., Mohapatra,T. and Singh,N.K.

Genomic sequence for Oryza sativa chromosome 11

Unpublished

2 (bases 1 to 174623)

Singh,N.K.

Direct Submission

Submitted (26-JAN-2002) Indian Initiative for Rice Genome

Sequencing, NRC Plant Biotechnology, Indian Agricultural Research

Institute, Pusa Road, New Delhi 110012, India

3 (bases 1 to 174623)

Sharma,T.R., McCombie,W., Mohapatra,T. and Singh,N.K.

Direct Submission

Submitted (12-JUN-2002) IIRGS, NRC on Plant Biotechnology, Indian

Agricultural Research Institute, LBS Centre, New Delhi, Delhi

110012, India

COMMENT

* NOTE: This is a 'working draft' sequence. It currently

* consists of 4 contigs. Gaps between the contigs

* are represented as runs of N. The order of the pieces

* is believed to be correct as given, however the sizes

* of the gaps between them are based on estimates that have

* been provided by the submitter.

* This sequence will be replaced

* by the finished sequence as soon as it is available and

* the accession number will be preserved.

* 1 3956: contig of 3956 bp in length

* 3957 4056: gap of unknown length

* 4057 8421: contig of 4365 bp in length

* 8422 8522: gap of unknown length

* 8522 73067: contig of 64546 bp in length

* 73068 73167: gap of unknown length

* 73168 174623: contig of 101456 bp in length.

* Location/Qualifiers

1. 174623

FEATURES

source

JOURNAL

COMMENT

Submitted (24-AUG-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
 On May 26, 2000 this sequence version replaced gi:6454072.
 All repeats were identified using RepeatMasker:
 Smit, A.F.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

----- Genome Center
 Center: Whitehead Institute/ MIT Center for Genome Research
 Center code: WIBR
 Web site: <http://www-seq.wi.mit.edu>
 Contact: sequence_submissions@genome.wi.mit.edu

----- Project Information
 Center project name: L1382
 Center clone name: 325_P15

----- Summary Statistics
 Sequencing vector: M13; M7815; 100% of reads
 Chemistry: Dye-terminator Big Dye; 100% of reads
 Assembly program: Phrap; version 0.960731
 Consensus quality: 141788 bases at least Q40
 Consensus quality: 163541 bases at least Q30
 Consensus quality: 174666 bases at least Q20
 Insert size: 181000; agarose-fp
 Insert size: 181005; sum-of-contigs
 Quality coverage: 4.2 in Q20 bases; agarose-fp
 Quality coverage: 4.2 in Q20 bases; sum-of-contigs

----- NOTE: This is a 'working draft' sequence. It currently
 * consists of 17 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

1 2035: contig of 2035 bp in length
 2036 2135: gap of 100 bp
 2136 6069: contig of 3934 bp in length
 6070 6169: gap of 100 bp
 6170 10460: contig of 4291 bp in length
 10461 10560: gap of 100 bp
 10561 14300: contig of 3740 bp in length
 14301 14400: gap of 100 bp
 14401 18480: contig of 4080 bp in length
 18481 18580: gap of 100 bp
 18581 23093: contig of 4513 bp in length
 23094 23193: gap of 100 bp
 23194 28213: contig of 5020 bp in length
 28214 28313: gap of 100 bp
 28314 32808: contig of 4495 bp in length
 32809 32908: gap of 100 bp
 32909 37000: contig of 4092 bp in length
 37001 37100: gap of 100 bp
 37101 45183: contig of 8083 bp in length
 45184 45283: gap of 100 bp
 45284 53421: contig of 8138 bp in length
 53422 53521: gap of 100 bp
 53522 64476: contig of 10955 bp in length
 64477 64576: gap of 100 bp
 64577 75144: contig of 10568 bp in length
 75145 75244: gap of 100 bp
 75245 93230: contig of 17986 bp in length
 93231 93330: gap of 100 bp
 93331 116951: contig of 23621 bp in length
 116952 117051: gap of 100 bp
 117052 143611: contig of 26560 bp in length
 143612 143711: gap of 100 bp
 143712 182695: contig of 38984 bp in length.
 Location/Qualifiers
 1. .182695
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="1"
 /map="1"

FEATURES
source

/clone="RP11-325P15"
 1. .2035
 /note="assembly_fragment"
 2136. .6069
 /note="assembly_fragment"
 6170. .10460
 /note="assembly_fragment"
 10561. .14300
 /note="assembly_fragment"
 14401. .18480
 /note="assembly_fragment"
 18581. .23093
 /note="assembly_fragment"
 23194. .28213
 /note="assembly_fragment"
 28314. .32808
 /note="assembly_fragment"
 clone_end:77
 vector_side:right
 32909. .37000
 /note="assembly_fragment"
 37101. .45183
 /note="assembly_fragment"
 45284. .53421
 /note="assembly_fragment"
 clone_end:SP6
 vector_side:left
 53522. .64476
 /note="assembly_fragment"
 64577. .75144
 /note="assembly_fragment"
 75245. .93230
 /note="assembly_fragment"
 93331. .116951
 /note="assembly_fragment"
 117052. .143611
 /note="assembly_fragment"
 143712. .182695
 /note="assembly_fragment"
 BASE COUNT 56372 a 36165 c 37419 g 51126 t 1613 others
 ORIGIN
 Query Match 85.5%; Score 18.8; DB 2; Length 182695;
 Best Local Similarity 90.9%; Pred.No.2.3e+02;
 Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 1 GACGTTGACACGGCGGTTCT 22
 |||||
 Db 19692 GACGTTGACACAGTGGTTCT 19671
 |||||
 RESULT 15
 NGU43735
 LOCUS
 DEFINITION
 NGU43735 2219 bp DNA linear BCT 30-JAN-1996
 Neisseria gonorrhoeae restriction/modification system, M.Ngov
 cytosine methylase (dcmE) gene, complete cds, and R.Ngov
 restriction endonuclease (dcrE) gene, partial cds.
 U43735
 U43735.1 GI:1165240
 ACCSSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Neisseria gonorrhoeae strain=WR302.
 Neisseria gonorrhoeae
 Bacteria; Proteobacteria; beta subdivision; Neisseriaceae;
 Neisseria.
 REFERENCE
 1 (bases 1 to 2219)
 AUTHORS
 TITLE
 JOURNAL
 MEDLINE
 PUBMED
 REFERENCE 2 (bases 1 to 2219)
 AUTHORS
 Stein,D.C., Gunn,J.S., Radlinska,M. and Piekarowicz,A.
 Restriction and modification systems of Neisseria gonorrhoeae
 Gene 157 (1-2), 19-22 (1995)
 9531562
 7607490

Mon Apr 21 10:30:32 2003

Direct Submission
 TITLE Submitted (20-DEC-1995) Daniel C. Stein, Microbiology, Univ.
 JOURNAL Maryland, College Park, MD 20742, USA
 FEATURES Location/Qualifiers
 source
 1. .2219
 /organism="Neisseria gonorrhoeae"
 /strain="WR302"
 /db_xref="taxon:485"
 /clone="pJM5"
 1. .241
 /gene="leuD"
 <1. .241
 /gene="leuD"
 /function="leucine biosynthesis"
 /codon_start=2
 /transl_table=11
 /product="isopropylmalate isomerase"
 /protein_id="AAA86267.1"
 /db_xref="GI:1165241"
 /translation="VDRPFKEVEANEGRSLSDLAEOQLTTPGGTFTFDITEHRKHC
 LNLGLDEIGLTLQHADKIKAFBEKRRSQSOWLFNG"
 408. .1679
 /gene="dcme"
 408. .1679
 /gene="dcme"
 /function="cytosine methylase"
 /codon_start=1
 /transl_table=11
 /product="M.NgoV DNA methylase"
 /protein_id="AAA86268.1"
 /db_xref="GI:1165242"
 /translation="MQQIKFIDLPSGMSGRKGFQACRQKQSVACCEVFTSRIKPAAL
 EVLKQNPYDVPYGDITKIETDIPDFDILLAGFPQAFSPACKRGLFEDTGTLEFFD
 VARILKAKPKGTLLENVEGLVTHDRKDSQKIGRTLVILEALGIVYVSKVLNA
 KEFGIPONRKRIYLTGSKSPDLSPFSPKNIILSGLPSTSPFIKLLKKFP
 PSELYGKSVKDKGKGNHSDIELKAVTEEEKQLNAILLKERKKGLOKIGIDW
 MDGMLTYAQISFYKHPDLONLDSITDKGYLVLEHPKQIGGQIKGQIKDESIFPKGYNI
 VSGKKSFEINKILDPNDVAPTLFAMDMEHLFVVDNGGLRLTLTGKGLRLFGYPDDYFP
 DIPKKDRCDLLGNTVAVPVINKAVSERLHL"
 1695. .2219
 /gene="dcrE"
 1695. .2219
 /gene="dcrE"
 /function="restriction enzyme"
 /codon_start=1
 /transl_table=11
 /product="R.NgoV restriction endonuclease"
 /protein_id="AAA86269.1"
 /db_xref="GI:1165243"
 /translation="MILTAQOIFDKLLDEKILSANGQIRFFLGVDVITVKKDVVG
 NIIQWLGWLRKEIEFDVSTNTQMPDPFNLKDRSRELLLEVAFNRNACPGFDIA
 DFKMYSDSFISPIGCGCLFNIGYMDMDNGNVTIKDLWLKKVQITRSMGWAINFKS
 KKAWCIRKARVFGTA"
 741 a 427 c 476 g 575 t
 BASE COUNT
 ORIGIN
 Query Match 83.6%; Score 18.4; DB 1; Length 2219;
 Best Local Similarity 95.0%; Pred. No. 5.2e+02;
 Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 GACGTTGACAAACGGCGGTTT 20
 Db 1503 GTCGTTGACAAACGGCGGTTT 1522

Search completed: April 18, 2003, 06:31:06
 Job time : 501 secs

GenCore version 5.1.4 p5 4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 18, 2003, 04:12:35 ; Search time 98 seconds
(without alignments)
505.551 Million cell updates/sec

Title: US-09-270-437D-14

Perfect score: 22
Sequence: 1 gacgtgacacgcgggtttct 22

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_101002.*

- 1: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.*
- 2: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.*
- 3: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.*
- 4: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.*
- 5: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.*
- 6: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.*
- 7: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.*
- 8: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.*
- 9: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.*
- 10: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.*
- 11: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.*
- 12: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.*
- 13: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.*
- 14: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.*
- 15: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.*
- 16: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.*
- 17: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.*
- 18: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.*
- 19: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.*
- 20: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.*
- 21: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.*
- 22: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.*
- 23: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.*
- 24: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	22	100.0	22	21	AAZ36160
2	22	100.0	424	21	AAH31011
3	22	100.0	2010	23	AAH70981
4	22	100.0	3283	21	AAZ36154
5	22	100.0	3412	21	AAZ36152
6	22	100.0	3667	23	AAH70982
7	22	100.0	3694	22	AAH26150
8	18.4	83.6	999	24	ABL87975
9	18.4	83.6	999	24	ABL87976

C 10	18.4	83.6	999	24	ABL87977	DNA polymerase III
C 11	18.4	83.6	999	24	ABL87978	DNA polymerase III
C 12	18.4	83.6	12893	21	AAH81733	N. meningitidis pa
C 13	18.4	83.6	102634	21	AAH81464	N. meningitidis pa
C 14	18.4	83.6	349980	21	AAH21608	Neisseria meningit
C 15	18.4	83.6	349980	21	AAH21609	Neisseria meningit
C 16	18.4	83.6	1437668	21	AAH81490	N. meningitidis B
C 17	17.2	78.2	140	22	ABA11525	Human nervous syst
C 18	17.2	78.2	404	21	AAH27439	Human secreted pro
C 19	17.2	78.2	7330	23	ABL06499	Drosophila melanog
C 20	17.2	78.2	10543	23	ABL06498	Drosophila melanog
C 21	17.2	78.2	349980	22	AAH41226	Pyrococcus abyssi
C 22	17.2	77.3	3378	23	ABL20008	Drosophila melanog
C 23	16.8	76.4	512	21	AAH12699	Aspergillus oryzae
C 24	16.8	76.4	969	21	AAH47512	Arabidopsis thalia
C 25	16.8	76.4	974	21	AAH34520	Arabidopsis thalia
C 26	16.2	73.6	264	21	AAH43435	Arabidopsis thalia
C 27	16.2	73.6	294	24	ABN28320	Arabidopsis thalia
C 28	16.2	73.6	525	20	AAH27120	Human ORFX polynuc
C 29	16.2	73.6	594	21	AAH50776	Calcineurin regula
C 30	16.2	73.6	683	21	AAH50785	Arabidopsis thalia
C 31	16.2	73.6	696	21	AAH34394	Arabidopsis thalia
C 32	16.2	73.6	883	21	AAH40090	Arabidopsis thalia
C 33	16.2	73.6	973	24	ABN21924	Arabidopsis thalia
C 34	16.2	73.6	1026	20	AAH89483	Human ORFX polynuc
C 35	16.2	73.6	1026	20	AAH31556	Bacillus lichenifo
C 36	16.2	73.6	1026	21	AAH52302	B. licheniformis p
C 37	16.2	73.6	1026	22	AAH86593	B. subtilis pectat
C 38	16.2	73.6	1035	24	ABK75359	DNA encoding the m
C 39	16.2	73.6	1157	18	AAH75430	Bacillus lichenifo
C 40	16.2	73.6	1390	24	AAH42599	Xanthomonas campe
C 41	16.2	73.6	6250	19	AAH58731	Human serine/threo
C 42	16.2	73.6	6250	21	AAH75306	Pyroclonitrin gene
C 43	16.2	73.6	8931	21	AAH75307	DNA sequence of Ps
C 44	16.2	73.6	56050	23	AAH59549	DNA sequence of Bu
C 45	16.2	73.6	56485	21	AAH81476	Propionibacterium
						N. meningitidis pa

ALIGNMENTS

RESULT 1
AAZ36160
ID AAZ36160 standard; DNA; 22 BP.

XX AAZ36160;

XX 11-FEB-2000 (first entry)

XX PCR primer for DNA encoding cancer associated antigen KOC-3.

XX Cancer associated antigen; KOC-1; cancer; vaccine; CT7; PCR primer; ss.

OS Synthetic.

OS Homo sapiens.

FN WO9954738-A1.

PD 28-OCT-1999.

PF 16-MAR-1999; 99WO-US05766.

XX 17-APR-1998; 98US-0061709.

XX (LUDW-) LUDWIG INST CANCER RES.

XX Chen Y, Gure A, Tsang S, Stockert E, Jager E, Knuth A, Old LJ;

XX WFI; 2000-013284/01.

XX Nucleotides representing cancer-associated genes, used to develop products for the diagnosis, monitoring and treatment of cancers

one of the 1079 sequences; (2) a recombinant host cell containing (I); (3) an isolated polypeptide (II) encoded by (I); (4) an antibody that specifically binds to (II); (5) a vector comprising (I); and (6) a method of detecting differentially expressed genes correlated with a cancerous state of a mammalian cell comprising detecting a gene product encoded by one of the 1079 sequences given in the specification. The polynucleotides are used to monitor patients having (or susceptible) to cancer to detect potentially malignant events at a molecular level before they are detectable at a gross morphological level. The polynucleotides are also useful for monitoring the efficacy of various therapies and preventive interventions. Polynucleotide probes based on the disclosed sequences are useful for chromosome mapping and detection of transcription levels. The 1079 polynucleotide sequences were derived from a human colon cancer cell line Kml2L4-A cDNA library.

Sequence 424 BP; 117 A; 106 C; 111 G; 90 T; 0 other;

Query Match 100.0%; Score 22; DB 21; Length 424;
Best Local Similarity 100.0%; Pred. No. 0.43; 0; Indels 0; Gaps 0;
Matches 22; Conservative 0; Mismatches 0;

QY 1 GACGTTGACAAACGGCGTTTCT 22
DB 240 GACGTTGACAAACGGCGTTTCT 219

RESULT 3
AA570981/c
ID AA570981 standard; cDNA; 2010 BP.
XX
AC AA570981;
XX
DT 13-FEB-2002 (first entry)
XX
DE DNA encoding novel human diagnostic protein #6785.
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
OS Homo sapiens.
XX
PN WO200175067-A2.
XX
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US08631.
XX
PR 31-MAR-2000; 2000US-0540217.
PR 23-AUG-2000; 2000US-0649167.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Drmanac RT, Liu C, Tang YT;
XX
DR WPI; 2001-639362/73.
DR P-PSDB; ABG06794.
XX
PT New isolated polynucleotide and encoded polypeptides, useful in diagnostics, forensics, gene mapping, identification of mutations responsible for genetic disorders or other traits and to assess biodiversity -
XX
PS Claim 1; SEQ ID No 6785; 103pp; English.
XX
CC The invention relates to isolated polynucleotide (I) and polypeptide (II) sequences. (I) is useful as hybridisation probes, polymerase chain reaction (PCR) primers, oligomers, and for chromosome mapping, and in recombinant production of (II). The polynucleotides are also used in diagnostics as expressed sequence tags for identifying expressed genes. (I) is useful in gene therapy techniques to restore normal activity of (II) or to treat disease states involving (II). (II) is useful for generating antibodies against it, detecting or quantitating a polypeptide in tissue, as molecular weight markers and as

PCR primers AA236159-60 were used to amplify a cancer associated antigen gene designated KOC-3. The specification also describes a cancer associated antigen designated CT7. The CT7 polynucleotide was isolated from SK-MEL-37 melanoma cells. The polypeptide has some homology with MAGE-10, limited to about 210 carboxy terminal amino acids. The amino terminal of the protein has a repetitive pattern, with repeats rich in serine, proline, glutamine and leucine, and an almost invariable core of the peptide given in AA43877. The CT7 polypeptide can be processed to peptides which provoke lysis by cytolytic T cells. The polynucleotides and polypeptides can be used for treating a cancerous condition and screening for or diagnosing cancerous conditions. The cancer associated antigens can be used as an immunogenic or vaccine composition with an adjuvant, e.g. a cytokine, a saponin, or granulocyte macrophage-colony stimulating factor (GM-CSF).

Sequence 22 BP; 4 A; 5 C; 7 G; 6 T; 0 other;

Query Match 100.0%; Score 22; DB 21; Length 22;
Best Local Similarity 100.0%; Pred. No. 0.36;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACGTTGACAAACGGCGTTTCT 22
DB 1 GACGTTGACAAACGGCGTTTCT 22

RESULT 2
AAH31011/c
ID AAH31011 standard; cDNA; 424 BP.
XX
AC AAH31011;
XX
DT 27-JUL-2001 (first entry)
XX
DE Human colon cancer cell line Kml2L4-A cDNA library derived sequence #945.
XX
KW Human; diagnosis; colon cancer; cancer; malignant; chromosome mapping;
KW detection; colon cancer cell line Kml2L4-A; ss.
XX
OS Homo sapiens.
XX
PN WO200018916-A2.
XX
PD 06-APR-2000.
XX
PF 23-SEP-1999; 99WO-US22226.
XX
PR 28-SEP-1998; 98US-0102161.
PR 28-SEP-1998; 98US-0102180.
PR 29-SEP-1998; 98US-0102380.
PR 08-OCT-1998; 98US-0103815.
PR 27-OCT-1998; 98US-0105977.
XX
PA (CHIR) CHIRON CORP.
PA (HYSE-) HYSEQ INC.
XX
PI Williams LT, Escobedo J, Innis MA, Garcia PD, Sudduth-Klinger J;
PI Reinhard C, Giese K, Randazzo F, Kennedy GC, Pot D, Kassam A;
PI Lamson G, Drmanac R, Crkvenjakov R, Dickinson M, Drmanac S, Labat I;
PI Leshkowitz D, Kita D, Garcia V, Jones LW, Stache-Crain B;
XX
DR WPI; 2000-293155/25.
XX
PT Polynucleotide library comprising 1079 defined sequences, useful in the form of an array to detect cancer or susceptibility to cancer -
XX
PS Claim 1; Page 454; 502pp; English.
XX
CC The present invention describes a library of polynucleotides comprising 1079 nucleotide sequences (given in AAH30067 to AAH31145). Also described are: (1) an isolated polynucleotide (I) having at least 90% identity to

CC a food supplement. (II) and its binding partners are useful in medical
 CC imaging of sites expressing (II). (I) and (II) are useful for treating
 CC disorders involving aberrant protein expression or biological activity.
 CC The polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. AAS64197-AAS94564 represent novel human
 CC diagnostic coding sequences of the invention.
 CC Note: The sequence data for this patent did not appear in the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.

SQ Sequence 2010 BP; 565 A; 552 C; 532 G; 361 T; 0 other;

Query Match 100.0%; Score 22; DB 23; Length 2010;
 Best Local Similarity 100.0%; Pred. No. 0.48; 0; Indels 0; Gaps 0;
 Matches 22; Conservative 0; Mismatches 0;

QY 1 GACGTTGACAAACGGCGTTTCT 22
 Db 421 GACGTTGACAAACGGCGTTTCT 400

RESULT 4

AAZ36154/c
 ID AAZ36154 standard; DNA; 3283 BP.

XX AAZ36154;

XX 11-FEB-2000 (first entry)

XX An alternative form of DNA encoding cancer associated antigen KOC-3.

XX Cancer associated antigen; KOC-3; cancer; vaccine; CT7; ss.

XX Homo sapiens.

XX WO9954738-A1.

XX 28-OCT-1999.

XX 16-MAR-1999; 99WO-US05766.

XX 17-APR-1998; 98US-0061709.

XX (LUDW-) LUDWIG INST CANCER RES.

XX Chen Y, Gure A, Tsang S, Stockert E, Jager E, Knuth A, Old LJ;
 WPI; 2000-013284/01.

XX Nucleotides representing cancer-associated genes, used to develop
 XX products for the diagnosis, monitoring and treatment of cancers.

XX Claim 55; Page 42-43; 44pp; English.

XX The present sequence represents an alternative form of a cancer
 XX associated antigen gene designated KOC-3. The specification also
 XX describes a cancer associated antigen designated CT7. The CT7
 XX polynucleotide has been isolated from SK-MEL-37 melanoma cells. The
 XX polypeptide has some homology with MAGE-10, limited to about 210 carboxy
 XX terminal amino acids. The amino terminal of the protein has a repetitive
 XX pattern, with repeats rich in serine, proline, glutamine and leucine,
 XX and an almost invariable core of the peptide given in AAY43877. The CT7
 XX polypeptide can be processed to peptides which provoke lysis by
 XX cytolytic T cells. The polynucleotides and polypeptides can be used for
 XX treating a cancerous condition and screening for or diagnosing cancerous
 XX conditions. The cancer associated antigens can be used as an immunogenic
 XX or vaccine composition with an adjuvant, e.g. a cytokine, a saponin, or
 XX granulocyte macrophage-colony stimulating factor (GM-CSF).

XX Sequence 3283 BP; 945 A; 833 C; 779 G; 725 T; 1 other;

Query Match 100.0%; Score 22; DB 21; Length 3283;
 Best Local Similarity 100.0%; Pred. No. 0.49; 0; Indels 0; Gaps 0;
 Matches 22; Conservative 0; Mismatches 0;

QY 1 GACGTTGACAAACGGCGTTTCT 22
 Db 444 GACGTTGACAAACGGCGTTTCT 423

RESULT 5

AAZ36152/c
 ID AAZ36152 standard; DNA; 3412 BP.

XX AAZ36152;

XX 11-FEB-2000 (first entry)

XX DNA encoding cancer associated antigen KOC-3.

XX Cancer associated antigen; KOC-3; cancer; vaccine; CT7; ss.

XX Homo sapiens.

XX WO9954738-A1.

XX 28-OCT-1999.

XX 16-MAR-1999; 99WO-US05766.

XX 17-APR-1998; 98US-0061709.

XX (LUDW-) LUDWIG INST CANCER RES.

XX Chen Y, Gure A, Tsang S, Stockert E, Jager E, Knuth A, Old LJ;
 WPI; 2000-013284/01.

XX Nucleotides representing cancer-associated genes, used to develop
 XX products for the diagnosis, monitoring and treatment of cancers.

XX Claim 55; Page 41; 44pp; English.

XX The present sequence represents a cancer associated antigen gene
 XX designated KOC-3. The specification also describes a cancer associated
 XX antigen designated CT7. The CT7 polynucleotide was isolated from
 XX SK-MEL-37 melanoma cells. The polypeptide has some homology with
 XX MAGE-10, limited to about 210 carboxy terminal amino acids. The amino
 XX terminal of the protein has a repetitive pattern, with repeats rich in
 XX serine, proline, glutamine and leucine, and an almost invariable core of
 XX the peptide given in AAY43877. The CT7 polypeptide can be processed to
 XX peptides which provoke lysis by cytolytic T cells. The polynucleotides
 XX and polypeptides can be used for treating a cancerous condition and
 XX screening for or diagnosing cancerous conditions. The cancer associated
 XX antigens can be used as an immunogenic or vaccine composition with an
 XX adjuvant, e.g. a cytokine, a saponin, or granulocyte macrophage-colony
 XX stimulating factor (GM-CSF).

XX Sequence 3412 BP; 970 A; 887 C; 804 G; 750 T; 1 other;

Query Match 100.0%; Score 22; DB 21; Length 3412;
 Best Local Similarity 100.0%; Pred. No. 0.49; 0; Indels 0; Gaps 0;
 Matches 22; Conservative 0; Mismatches 0;

QY 1 GACGTTGACAAACGGCGTTTCT 22
 Db 444 GACGTTGACAAACGGCGTTTCT 423

RESULT 6

AAS70982/c
 ID AAS70982 standard; cDNA; 3667 BP.

XX

[illegible]

OS Neisseria meningitidis.
 PN WO200206532-A1.
 PD 24-JAN-2002.
 XX 16-JUL-2001; 2001WO-US222395.
 PF 14-JUL-2000; 2000US-218246P.
 PR 28-MAR-2001; 2001US-0818780.
 XX (REPL-) REPLIDYNE INC.
 PA Bullard JJ, Janjic N, McHenry CS;
 PI WPI; 2002-164785/21.
 XX P-PSDB; ABB84744.
 DR Nucleic acid encoding a DNA polymerase III holoenzyme delta subunit
 XX from bacteria, useful for screening agents that modulate the subunit
 PT activity which is useful in the treatment of bacterial infections e.g.
 PT S. pyogenes and S. aureus -
 XX Disclosure; Page 282; 500pp; English.
 PS The present invention describes nucleic acid sequences encoding a DNA
 XX polymerase III holoenzyme delta subunit (I) from bacteria. (I) has
 CC antibacterial; tuberculostatic; antileprotic. Methods from the present
 CC invention can be used for screening for bacterial DNA polymerase
 CC holoenzyme delta subunit proteins and agents that modulate their
 CC activity. The agents are useful in the treatment of bacterial
 CC infections e.g. S. pyogenes, S. aureus, S. pneumoniae, Mycoplasma,
 CC Yersinia, Corynebacterium, Salmonella, Mycobacterium tuberculosis or
 CC M. leprae. The invention provides a convenient means of identifying
 CC compounds which modulate DNA replication in bacteria and therefore
 CC provide antibacterial targets, and which are also useful for
 CC amplification of DNA. ABL87935 to ABL88071 and ABB84724 to ABB84816
 CC represent sequences used in the exemplification of the present
 CC invention.
 XX Sequence 999 BP; 219 A; 262 C; 328 G; 190 T; 0 other;
 SQ Query Match 83.6%; Score 18.4; DB 24; Length 999;
 Best Local Similarity 95.0%; Pred. No. 25;
 Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 GACGTTGACACGCGGTTT 20
 DB 621 GACGTTGGCAACGCGGTTT 602
 RESULT 9
 ABL87976/C
 ID ABL87976 standard; DNA; 999 BP.
 XX ABL87976;
 AC 16-MAY-2002 (first entry)
 XX DNA polymerase III holoenzyme delta subunit related DNA SEQ ID NO:62.
 DT DNA polymerase III holoenzyme delta subunit; DNA polymerase III;
 DE DnaX; hola; hola; antibacterial; tuberculostatic; antileprotic;
 XX bacterial infection; DNA replication modulation; gene; ds.
 KW Neisseria meningitidis.
 XX WO200206532-A1.
 XX 24-JAN-2002.
 PD 16-JUL-2001; 2001WO-US222395.

XX 14-JUL-2000; 2000US-218246P.
 PR 28-MAR-2001; 2001US-0818780.
 XX (REPL-) REPLIDYNE INC.
 PA Bullard JJ, Janjic N, McHenry CS;
 PI WPI; 2002-164785/21.
 XX P-PSDB; ABB84744.
 DR Nucleic acid encoding a DNA polymerase III holoenzyme delta subunit
 XX from bacteria, useful for screening agents that modulate the subunit
 PT activity which is useful in the treatment of bacterial infections e.g.
 PT S. pyogenes and S. aureus -
 XX Disclosure; Page 283-284; 500pp; English.
 PS The present invention describes nucleic acid sequences encoding a DNA
 XX polymerase III holoenzyme delta subunit (I) from bacteria. (I) has
 CC antibacterial; tuberculostatic; antileprotic. Methods from the present
 CC invention can be used for screening for bacterial DNA polymerase
 CC holoenzyme delta subunit proteins and agents that modulate their
 CC activity. The agents are useful in the treatment of bacterial
 CC infections e.g. S. pyogenes, S. aureus, S. pneumoniae, Mycoplasma,
 CC Yersinia, Corynebacterium, Salmonella, Mycobacterium tuberculosis or
 CC M. leprae. The invention provides a convenient means of identifying
 CC compounds which modulate DNA replication in bacteria and therefore
 CC provide antibacterial targets, and which are also useful for
 CC amplification of DNA. ABL87935 to ABL88071 and ABB84724 to ABB84816
 CC represent sequences used in the exemplification of the present
 CC invention.
 XX Sequence 999 BP; 219 A; 263 C; 327 G; 190 T; 0 other;
 SQ Query Match 83.6%; Score 18.4; DB 24; Length 999;
 Best Local Similarity 95.0%; Pred. No. 25;
 Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 GACGTTGACACGCGGTTT 20
 DB 621 GACGTTGGCAACGCGGTTT 602
 RESULT 10
 ABL87977/C
 ID ABL87977 standard; DNA; 999 BP.
 XX ABL87977;
 AC 16-MAY-2002 (first entry)
 XX DNA polymerase III holoenzyme delta subunit related DNA SEQ ID NO:64.
 DT DNA polymerase III holoenzyme delta subunit; DNA polymerase III;
 DE DnaX; hola; hola; antibacterial; tuberculostatic; antileprotic;
 XX bacterial infection; DNA replication modulation; gene; ds.
 KW Neisseria meningitidis.
 XX WO200206532-A1.
 XX 24-JAN-2002.
 PD 16-JUL-2001; 2001WO-US222395.
 PF 14-JUL-2000; 2000US-218246P.
 PR 28-MAR-2001; 2001US-0818780.
 XX (REPL-) REPLIDYNE INC.
 PA Bullard JJ, Janjic N, McHenry CS;
 PI 14-JUL-2000; 2000US-218246P.
 PR 28-MAR-2001; 2001US-0818780.
 XX (REPL-) REPLIDYNE INC.

DR WPI; 2002-164785/21.
 DR P-PSDB; ABB84745.
 XX
 PT Nucleic acid encoding a DNA polymerase III holoenzyme delta subunit
 PT from bacteria, useful for screening agents that modulate the subunit
 PT activity which is useful in the treatment of bacterial infections e.g.
 PT S. pyogenes and S. aureus -
 XX
 XX
 XX Disclosure; Fig 4U; 500pp; English.
 XX
 CC The present invention describes nucleic acid sequences encoding a DNA
 CC polymerase III holoenzyme delta subunit (I) from bacteria. (I) has
 CC antibacterial; tuberculostatic; antileprotic. Methods from the present
 CC invention can be used for screening for bacterial DNA polymerase
 CC holoenzyme delta subunit proteins and agents that modulate their
 CC activity. The agents are useful in the treatment of bacterial
 CC infections e.g. S. pyogenes, S. aureus, S. pneumoniae, Mycoplasma,
 CC Yersinia, Corynebacterium, Salmonella, Mycobacterium tuberculosis or
 CC M. leprae. The invention provides a convenient means of identifying
 CC compounds which modulate DNA replication in bacteria and therefore
 CC provide antibacterial targets, and which are also useful for
 CC amplification of DNA. ABL87935 to ABL88071 and ABB84724 to ABB84816
 CC represent sequences used in the exemplification of the present
 CC invention.
 XX
 SQ Sequence 999 BP; 222 A; 260 C; 327 G; 190 T; 0 other;
 Query Match 83.6%; Score 18.4; DB 24; Length 999;
 Best Local Similarity 95.0%; Pred. No. 25;
 Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 GACGTTGACACGCGGTTT 20
 Db 621 GACGTTGACACGCGGTTT 602
 RESULT 11
 ABL87978/C
 ID ABL87978 standard; DNA; 999 BP.
 AC ABL87978;
 XX
 XX 16-MAY-2002 (first entry)
 DT
 DE DNA polymerase III holoenzyme delta subunit related DNA SEQ ID NO:65.
 XX
 KW DNA polymerase III holoenzyme delta subunit; DNA polymerase III;
 KW DnaX; hoiA; hoiB; antibacterial; tuberculostatic; antileprotic;
 KW bacterial infection; DNA replication modulation; gene; ds.
 XX
 OS Neisseria meningitidis.
 XX
 PN WO200206532-A1.
 PD
 XX 24-JAN-2002.
 PF
 XX 16-JUL-2001; 2001WO-US22395.
 XX
 PR 14-JUL-2000; 2000US-218246P.
 XX
 PR 28-MAR-2001; 2001US-0818780.
 XX
 PA (REPL-) REPLIDYNE INC.
 XX
 XX Bullard JJ, Janjic N, McHenry CS;
 PI
 XX WPI; 2002-164785/21.
 DR
 DR P-PSDB; ABB84745.
 XX
 XX Nucleic acid encoding a DNA polymerase III holoenzyme delta subunit
 XX from bacteria, useful for screening agents that modulate the subunit
 XX activity which is useful in the treatment of bacterial infections e.g.
 XX S. pyogenes and S. aureus -

PS Disclosure; Page 287-289; 500pp; English.
 XX
 CC The present invention describes nucleic acid sequences encoding a DNA
 CC polymerase III holoenzyme delta subunit (I) from bacteria. (I) has
 CC antibacterial; tuberculostatic; antileprotic. Methods from the present
 CC invention can be used for screening for bacterial DNA polymerase
 CC holoenzyme delta subunit proteins and agents that modulate their
 CC activity. The agents are useful in the treatment of bacterial
 CC infections e.g. S. pyogenes, S. aureus, S. pneumoniae, Mycoplasma,
 CC Yersinia, Corynebacterium, Salmonella, Mycobacterium tuberculosis or
 CC M. leprae. The invention provides a convenient means of identifying
 CC compounds which modulate DNA replication in bacteria and therefore
 CC provide antibacterial targets, and which are also useful for
 CC amplification of DNA. ABL87935 to ABL88071 and ABB84724 to ABB84816
 CC represent sequences used in the exemplification of the present
 CC invention.
 XX
 SQ Sequence 999 BP; 222 A; 260 C; 327 G; 190 T; 0 other;
 Query Match 83.6%; Score 18.4; DB 24; Length 999;
 Best Local Similarity 95.0%; Pred. No. 25;
 Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 GACGTTGACACGCGGTTT 20
 Db 621 GACGTTGACACGCGGTTT 602
 RESULT 12
 AAA81733/C
 ID AAA81733 standard; DNA; 12893 BP.
 AC AAA81733;
 XX
 XX 04-DEC-2000 (first entry)
 DT
 DE N. meningitidis partial DNA sequence gnm_280 SEQ ID NO:280.
 XX
 KW Neisseria meningitidis; Neisseria gonorrhoeae; genome; immunogenic;
 KW antigen; vaccine; diagnosis; infection; antibacterial; identification;
 KW Meningococcus B; MenB; ds.
 XX
 OS Neisseria meningitidis.
 XX
 PN WO200022430-A2.
 XX
 PD 20-APR-2000.
 XX
 PF 08-OCT-1999; 99WO-US23573.
 XX
 PR 09-OCT-1998; 98US-0103794.
 PR 30-APR-1999; 99US-0132068.
 XX
 PA (CHIR) CHIRON CORP.
 XX
 XX Frazer CM, Hickey E, Peterson J, Tettelin H, Venter JC;
 PI Masighani V, Galeotti C, Mora M, Ratti G, Scarselli M, Scariato V;
 PI Rappuoli R, Pizza M;
 XX
 DR WPI; 2000-318079/27.
 XX
 XX Isolated nucleotide sequences of Neisseria meningitidis which can be
 XX used in the diagnosis and treatment of N. meningitidis infection and
 XX other Neisserial infections, for example, N.gonorrhoea -
 PS
 PS Claim 7; Page 1561-1565; 1760pp; English.
 XX
 CC The present invention describes methods of obtaining immunogenic
 CC proteins from Neisseria genomic sequences. AAA81453 to AAA82414
 CC represent specifically claimed Neisseria meningitidis genomic DNA
 CC sequences; AAA81260 to AAA81303 and AAB25620 to AAB25663 represent
 CC Neisseria DNA sequences and their corresponding proteins; AAA81254 to
 CC AAA81259 and AAA81304 to AAA81321 represent PCR primers used in the


```

XX CC The present invention describes the full length genome of
CC Neisseria meningitidis B (NMB). The sequences in AAF21544 and AAF21607
CC to AAF21613 represent fragments of the NMB genomic sequence, as the
CC sequence was too long to go in a record on its own it was split into 8
CC sequences which overlap each other at the beginning and end of each
CC sequence by 49980 bp (i.e. the last 49980 bp of AAF21544 is repeated at
CC the beginning of AAF21607, the last 49980 bp of AAF21607 are repeated at
CC the beginning of AAF21608, and so on). AAF21545 to AAF21588 encode the
CC Neisseria proteins given in AAB58550 to AAB58593, and AAF21589 to
CC AAF21606 represent PCR primers which are used in the exemplification of
CC the present invention. The NMB genome and fragments from it have
CC antibacterial activity, and can be used in vaccines and gene therapy.
CC Neisseria nucleic acids, proteins and/or antibodies which binds to the
CC proteins can be used in compositions for treating or preventing infection
CC due to Neisserial bacteria or as a diagnostic reagent for detecting the
CC presence of Neisserial bacteria or of antibodies raised to Neisserial
CC bacteria. Computers, computer memory, computer storage medium or computer
CC databases can be used in a search to identify open reading frames (ORFs)
CC or coding sequences within the NMB genome. The DNA sequences provide
CC further opportunities to find antigenic or immunogenic proteins which are
CC more effective in vaccines than the outer membrane proteins currently
CC used.
XX CC
SQ Sequence 349980 BP; 82523 A; 82940 C; 96712 G; 87805 T; 0 other;

Query Match      83.6%; Score 18.4; DB 21; Length 349980;
Best Local Similarity 95.0%; Pred. No. 36;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GACGTTGACACGCGGGTTT 20
    |||||
Db 140350 GACGTTGACACGCGGGTTT 140331

RESULT 15
AAF21609/c
ID AAF21609 standard; DNA; 349980 BP.
XX CC
AC AAF21609;
XX CC
DT 13-MAR-2001 (first entry)
XX CC
DE Neisseria meningitidis B nucleotide sequence SEQ ID NO:110.
XX CC
KW Neisseria meningitidis; Neisseria gonorrhoeae; immunogenic; vaccine;
KW diagnosis; antigen; detection; infection; gene therapy; antibacterial;
KW ds.
XX CC
OS Neisseria meningitidis.
XX CC
PN WO200066791-A1.
XX CC
PD 09-NOV-2000.
XX CC
PF 08-MAR-2000; 2000WO-US05928.
XX CC
PR 30-APR-1999; 99US-0132068.
PR 08-OCT-1999; 99WO-US23573.
PR 28-FEB-2000; 2000GB-0004695.
XX CC
PA (CHIR ) CHIRON CORP.
PA (GENO-) INST GENOMIC RES.
XX CC
PI Piza M, Hickey B, Peterson J, Tettelin H, Venter JC, Massignani V;
PI Galeotti C, Mora M, Ratti G, Scarselli M, Scarlato V, Rappuoli R;
PI Frazer CM, Grandi G;
XX CC
DR WPI; 2000-647603/62.
XX CC
PT Neisseria meningitidis B full length genome sequence and open reading
XX frames are used to detect, treat and prevent Neisserial infections -

```

PS Claim 7; Appendix A; 692pp; English.

XX CC The present invention describes the full length genome of
CC Neisseria meningitidis B (NMB). The sequences in AAF21544 and AAF21607
CC to AAF21613 represent fragments of the NMB genomic sequence, as the
CC sequence was too long to go in a record on its own it was split into 8
CC sequences which overlap each other at the beginning and end of each
CC sequence by 49980 bp (i.e. the last 49980 bp of AAF21544 is repeated at
CC the beginning of AAF21607, the last 49980 bp of AAF21607 are repeated at
CC the beginning of AAF21608, and so on). AAF21545 to AAF21588 encode the
CC Neisseria proteins given in AAB58550 to AAB58593, and AAF21589 to
CC AAF21606 represent PCR primers which are used in the exemplification of
CC the present invention. The NMB genome and fragments from it have
CC antibacterial activity, and can be used in vaccines and gene therapy.
CC Neisseria nucleic acids, proteins and/or antibodies which binds to the
CC proteins can be used in compositions for treating or preventing infection
CC due to Neisserial bacteria or as a diagnostic reagent for detecting the
CC presence of Neisserial bacteria or of antibodies raised to Neisserial
CC bacteria. Computers, computer memory, computer storage medium or computer
CC databases can be used in a search to identify open reading frames (ORFs)
CC or coding sequences within the NMB genome. The DNA sequences provide
CC further opportunities to find antigenic or immunogenic proteins which are
CC more effective in vaccines than the outer membrane proteins currently
CC used.

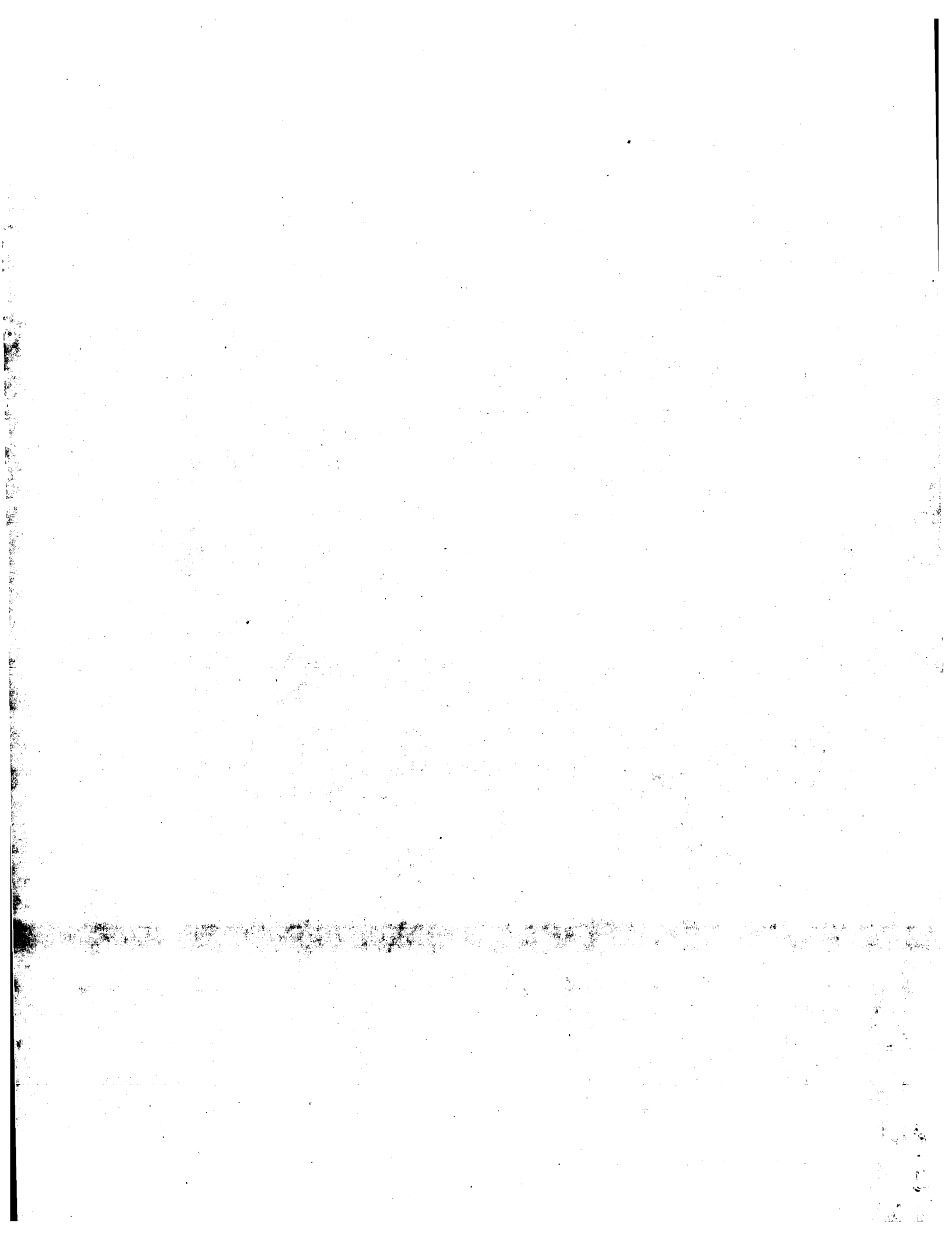
XX CC
SQ Sequence 349980 BP; 81351 A; 86755 C; 95584 G; 86290 T; 0 other;

Query Match 83.6%; Score 18.4; DB 21; Length 349980;
Best Local Similarity 95.0%; Pred. No. 36;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GACGTTGACACGCGGGTTT 20

Db 147755 GTCGTTGACACGCGGGTTT 147736

Search completed: April 18, 2003, 05:48:05
Job time : 251 secs



GenCore version 5.1.4.p5 4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 18, 2003, 04:53:41 ; Search time 21.5 Seconds
(without alignments)
313.809 Million cell updates/sec

Title: US-09-270-437D-14
Perfect score: 22
Sequence: 1 gacgttgacaacggcggtttct 22

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA.*
1: /cgm2_6/ptodata/1/ina/5A_COMB.seq.*
2: /cgm2_6/ptodata/1/ina/5B_COMB.seq.*
3: /cgm2_6/ptodata/1/ina/6A_COMB.seq.*
4: /cgm2_6/ptodata/1/ina/6B_COMB.seq.*
5: /cgm2_6/ptodata/1/ina/PCTUS_COMB.seq.*
6: /cgm2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	22	100.0	3283	4	US-09-061-709-8
C 2	22	100.0	3412	4	US-09-061-709-6
C 3	16.2	73.6	524	1	US-08-328-322-16
C 4	16.2	73.6	812	1	US-08-328-322-13
C 5	16.2	73.6	1026	3	US-09-073-684-1
C 6	16.2	73.6	1026	4	US-09-198-956-7
C 7	16.2	73.6	1026	4	US-09-198-955A-3
C 8	16.2	73.6	1026	4	US-09-395-858A-11
C 9	16.2	73.6	1026	4	US-09-546-500-1
C 10	16.2	73.6	1026	4	US-09-546-762-1
C 11	16.2	73.6	1026	4	US-09-694-531-3
C 12	16.2	73.6	1026	4	US-09-670-141-7
C 13	16.2	73.6	6250	1	US-08-729-214-23
C 14	16.2	73.6	6250	3	US-09-028-934-23
C 15	16.2	73.6	8931	3	US-09-028-934-28
C 16	15.8	71.8	537	4	US-08-836-500A-3
C 17	15.8	71.8	1007	4	US-08-836-500A-1
C 18	15.8	71.8	1008	3	US-08-721-979A-13
C 19	15.8	71.8	1008	4	US-09-654-289-13
C 20	15.6	70.9	831	4	US-08-998-416-290
C 21	15.6	70.9	1276	4	US-08-981-957D-12
C 22	15.6	70.9	3579	1	US-08-674-168-15
C 23	15.6	70.9	3579	3	US-08-985-908-18
C 24	15.6	70.9	3579	3	US-08-852-730-3
C 25	15.6	70.9	3579	4	US-08-985-916-10
C 26	15.6	70.9	4403765	4	US-09-103-840A-2
C 27	15.2	69.1	5134	2	US-08-635-121-1

C 28	15.2	69.1	6718	2	US-08-962-284-1	Sequence 1, Appli
C 29	15.2	69.1	10718	3	US-08-325-426B-1	Sequence 1, Appli
C 30	15.2	69.1	20235	1	US-07-642-734C-3	Sequence 3, Appli
C 31	15.2	69.1	20235	3	US-08-439-009A-3	Sequence 3, Appli
C 32	15	68.2	50	4	US-09-237-712-92	Sequence 92, Appli
C 33	15	68.2	397	4	US-09-237-712-1	Sequence 1, Appli
C 34	14.8	67.3	111282	4	US-09-754-250-3	Sequence 2, Appli
C 35	14.8	67.3	4403765	4	US-09-103-840A-2	Sequence 1, Appli
C 36	14.8	67.3	4411529	4	US-09-103-840A-1	Sequence 1, Appli
C 37	14.6	66.4	341	2	US-08-975-316-39	Sequence 39, Appli
C 38	14.6	66.4	341	4	US-09-615-192A-19	Sequence 39, Appli
C 39	14.6	66.4	409	2	US-08-975-316-41	Sequence 41, Appli
C 40	14.6	66.4	409	4	US-09-615-192A-41	Sequence 118, App
C 41	14.6	66.4	538	4	US-08-169-948B-15	Sequence 15, Appl
C 42	14.6	66.4	1155	2	US-08-448-873-15	Sequence 15, Appl
C 43	14.6	66.4	1155	2	US-08-382-452D-15	Sequence 15, Appl
C 44	14.6	66.4	1155	4	US-08-952-457-1	Sequence 1, Appli
C 45	14.6	66.4	1401	4		

ALIGNMENTS

RESULT 1

US-09-061-709-8/c
; Sequence 8, Application US/09061709B
; Patent No. 6297364
; GENERAL INFORMATION:
; APPLICANT: Chen, Yao-Tseng
; APPLICANT: Gure, Ali
; APPLICANT: Teang, Solam
; APPLICANT: Stockert, Elisabeth
; APPLICANT: Jager, Elke
; APPLICANT: Knuth, Alexander
; APPLICANT: Old, Lloyd J.
; TITLE OF INVENTION: Isolated Nucleic Acid Molecules Encoding Cancer Associated
; FILE OF INVENTION: Antigen, The Antigens Per Se, And Uses Thereof
; CURRENT APPLICATION NUMBER: US/09/061.709B
; CURRENT FILING DATE: 1998-04-17
; NUMBER OF SEQ ID NOS: 8
; SEQ ID NO 8
; LENGTH: 3283
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
US-09-061-709-8

Query Match 100.0%; Score 22; DB 4; Length 3283;
Best Local Similarity 100.0%; Pred. No. 0.062;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GACGTTGACAACGGCGGTTTCT 22
Db 444 GACGTTGACAACGGCGGTTTCT 423

RESULT 2

US-09-061-709-6/c
; Sequence 6, Application US/09061709B
; Patent No. 6297364
; GENERAL INFORMATION:
; APPLICANT: Chen, Yao-Tseng
; APPLICANT: Gure, Ali
; APPLICANT: Teang, Solam
; APPLICANT: Stockert, Elisabeth
; APPLICANT: Jager, Elke
; APPLICANT: Knuth, Alexander
; APPLICANT: Old, Lloyd J.
; TITLE OF INVENTION: Isolated Nucleic Acid Molecules Encoding Cancer Associated
; FILE OF INVENTION: Antigen, The Antigens Per Se, And Uses Thereof
; CURRENT APPLICATION NUMBER: US/09/061.709B

Mon Apr 21 10:30:38 2003

Db 307 GACATTGACACGAGCGGTTTC 327

RESULT 4

US-08-328-322-13

Sequence 13, Application US/08328322

Patent No. 5723436

GENERAL INFORMATION:

APPLICANT: Huang, Laiqiang

APPLICANT: Cyert, Martha S.

TITLE OF INVENTION: Calcineurin Interacting Protein Compositions

TITLE OF INVENTION: and Methods

NUMBER OF SEQUENCES: 23

CORRESPONDENCE ADDRESS:

ADDRESSEE: Dehlinger & Associates

STREET: 350 Cambridge Avenue, Suite 250

CITY: Palo Alto

STATE: CA

COUNTRY: USA

ZIP: 94306

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent in Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/328,322

FILING DATE: 24-OCT-1994

CLASSIFICATION: 530

ATTORNEY/AGENT INFORMATION:

NAME: Sholtz, Charles K.

REGISTRATION NUMBER: P38,615

REFERENCE/DOCKET NUMBER: 8600-0151.10

TELECOMMUNICATION INFORMATION:

TELEPHONE: (415) 324-0880

TELEFAX: (415) 324-0960

INFORMATION FOR SEQ ID NO: 13:

SEQUENCE CHARACTERISTICS:

LENGTH: 812 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: unknown

MOLECULE TYPE: DNA (genomic)

HYPOTHETICAL: NO

ANTI-SENSE: NO

ORIGINAL SOURCE:

INDIVIDUAL ISOLATE: genomic DNA fragment containing full

INDIVIDUAL ISOLATE: CNB1 coding sequence

FEATURE:

NAME/KEY: CDS

LOCATION: 54..104

NAME/KEY: CDS

LOCATION: 181..652

US-08-328-322-13

Query Match 73.6%; Score 16.2; DB 1; Length 812;

Best Local Similarity 85.7%; Pred. No. 35;

Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 GACGTTGACACGCGGTTTC 21

Db 436 GACATTGACACGAGCGGTTTC 456

RESULT 5

US-09-073-684-1

Sequence 1, Application US/09073684

Patent No. 6124127

GENERAL INFORMATION:

APPLICANT: Andersen, Lene No. 6124127boe

APPLICANT: Sch lein, Martin

APPLICANT: Lange, Niels Erik Krebs

;; TITLE OF INVENTION: No. 6124127el Pectate Lyases
;; NUMBER OF SEQUENCES: 10
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: No. 6124127o No. 6124127disk of No. 6124127th America, Inc.
;; STREET: 405 Lexington Avenue
;; CITY: New York
;; STATE: NY
;; COUNTRY: USA
;; ZIP: 10174
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Diskette
;; COMPUTER: IBM Compatible
;; OPERATING SYSTEM: DOS
;; SOFTWARE: FastSEQ for Windows Version 2.0
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/09/073,684
;; FILING DATE: 6-MAY-1998
;; CLASSIFICATION: 0506
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Carol E. Rozek
;; REGISTRATION NUMBER: 36,993
;; REFERENCE/DOCKET NUMBER: 5543.200-US
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 212-867-0123
;; TELEFAX: 212-878-9655
;; INFORMATION FOR SEQ ID NO: 1:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 1026 base pairs
;; TYPE: nucleic acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
US-09-073-684-1

Query Match 73.6%; Score 16.2; DB 3; Length 1026;
Best Local Similarity 85.7%; Pred. No. 36;
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 GACGTTGACACGCGGTTTC 21
Db 51 GTCATTGACAGCGCGGTTTC 71

RESULT 6
US-09-198-956-7
; Sequence 7, Application US/09198956
; Patent No. 6165769

;; GENERAL INFORMATION:
;; APPLICANT: Andersen, Lene N.
;; APPLICANT: Schulein, Martin
;; APPLICANT: Lange, Niels Erik K.
;; APPLICANT: Bjornvad, Mads E.
;; APPLICANT: Schnorr, Kirk
;; TITLE OF INVENTION: Pectin Degrading Enzymes From Bacillus
;; FILE REFERENCE: 5377.200-US
;; CURRENT APPLICATION NUMBER: US/09/198,956
;; CURRENT FILING DATE: 1998-11-24
;; EARLIER APPLICATION NUMBER: 1344/97
;; EARLIER FILING DATE: 1997-11-24
;; EARLIER APPLICATION NUMBER: 60/067,240
;; EARLIER FILING DATE: 1997-12-02
;; NUMBER OF SEQ ID NOS: 26
;; SOFTWARE: FastSEQ for Windows Version 3.0
;; SEQ ID NO 7
;; LENGTH: 1026
;; TYPE: DNA
;; ORGANISM: Bacillus licheniformis
US-09-198-956-7

Query Match 73.6%; Score 16.2; DB 4; Length 1026;
Best Local Similarity 85.7%; Pred. No. 36;
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 GACGTTGACACGCGGTTTC 21
Db 51 GTCATTGACAGCGCGGTTTC 71

RESULT 7
US-09-198-955A-3
; Sequence 3, Application US/09198955A
; Patent No. 6187580
;; GENERAL INFORMATION:
;; APPLICANT: Andersen, Lene N.
;; APPLICANT: Schulein, Martin
;; APPLICANT: Lange, Niels E.
;; APPLICANT: Bjornvad, Mads E.
;; APPLICANT: Moller, Soren
;; APPLICANT: Glad, Sanne O. S.
;; APPLICANT: Kauppinen, Markus S.
;; APPLICANT: Schnorr, Kirk
;; APPLICANT: Kongsbak, Lars
;; TITLE OF INVENTION: No. 6187580el Pectate Lyases
;; FILE REFERENCE: 5378.200-US
;; CURRENT APPLICATION NUMBER: US/09/198,955A
;; CURRENT FILING DATE: 1998-11-24
;; PRIOR APPLICATION NUMBER: 1343/97
;; PRIOR FILING DATE: 1997-11-24
;; PRIOR APPLICATION NUMBER: 1344/97
;; PRIOR FILING DATE: 1997-11-24
;; PRIOR APPLICATION NUMBER: 60/067,249
;; PRIOR FILING DATE: 1997-12-02
;; PRIOR APPLICATION NUMBER: 60/067,240
;; PRIOR FILING DATE: 1997-12-02
;; PRIOR APPLICATION NUMBER: 09/073,684
;; PRIOR FILING DATE: 1998-05-06
;; PRIOR APPLICATION NUMBER: 09/184,217
;; PRIOR FILING DATE: 1998-11-02
;; NUMBER OF SEQ ID NOS: 32
;; SOFTWARE: FastSEQ for Windows Version 4.0
;; SEQ ID NO 3
;; LENGTH: 1026
;; TYPE: DNA
;; ORGANISM: Bacillus licheniformis - ATCC 14580
US-09-198-955A-3

Query Match 73.6%; Score 16.2; DB 4; Length 1026;
Best Local Similarity 85.7%; Pred. No. 36;
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 GACGTTGACACGCGGTTTC 21
Db 51 GTCATTGACAGCGCGGTTTC 71

RESULT 8
US-09-395-858A-11
; Sequence 11, Application US/09395858A
; Patent No. 6242014
;; GENERAL INFORMATION:
;; APPLICANT: Feng Xu
;; TITLE OF INVENTION: Methods For Using Pectate Lyases In
;; FILE REFERENCE: 5670.200-US
;; CURRENT APPLICATION NUMBER: US/09/395,858A
;; CURRENT FILING DATE: 1999-09-14
;; PRIOR APPLICATION NUMBER: 09/156,298
;; PRIOR FILING DATE: 1998-09-17
;; NUMBER OF SEQ ID NOS: 12
;; SOFTWARE: FastSEQ for Windows Version 3.0
;; SEQ ID NO 11
;; LENGTH: 1026
;; TYPE: DNA
;; ORGANISM: Bacillus licheniformis
US-09-395-858A-11

RESULT 11
US-09-694-531-3
; Sequence 3, Application US/09694531
; Patent No. 636843
; GENERAL INFORMATION:
; APPLICANT: Andersen, Lene N.
; APPLICANT: Schulein, Martin
; APPLICANT: Lange, Niels E.
; APPLICANT: Bjornvad, Mads E.
; APPLICANT: Moller, Soren
; APPLICANT: Glad, Sanne O. S.
; APPLICANT: Kauppinen, Markus S.
; APPLICANT: Schnorr, Kirk
; APPLICANT: Kongsbak, Lars
; TITLE OF INVENTION: No 6368843el Pectate Lyases
; FILE REFERENCE: 5378-200-US
; CURRENT APPLICATION NUMBER: US/09/694,531
; CURRENT FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 09/198,955
; PRIOR FILING DATE: 1998-11-24
; PRIOR APPLICATION NUMBER: 1343/97
; PRIOR FILING DATE: 1997-11-24
; PRIOR APPLICATION NUMBER: 1344/97
; PRIOR FILING DATE: 1997-11-24
; PRIOR APPLICATION NUMBER: 60/067,249
; PRIOR FILING DATE: 1997-12-02
; PRIOR APPLICATION NUMBER: 60/067,240
; PRIOR FILING DATE: 1997-12-02
; PRIOR APPLICATION NUMBER: 09/073,684
; PRIOR FILING DATE: 1998-05-06
; PRIOR APPLICATION NUMBER: 09/184,217

;; PRIOR FILING DATE: 1998-11-02
;; NUMBER OF SEQ ID NOS: 32
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 3
;; LENGTH: 1026
;; TYPE: DNA
;; ORGANISM: Bacillus licheniformis - ATCC 14580
US-09-694-531-3

Query Match 73.6%; Score 16.2; DB 4; Length 1026;
Best Local Similarity 85.7%; Pred. No. 36;
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 GACGTTGACACGGCGGTTTC 21
Db 51 GTCATTGACACGGCGGTTTC 71

RESULT 12

US-09-670-141-7
; Sequence 7, Application US/09670141
; Patent No. 6429000
; GENERAL INFORMATION:
; APPLICANT: Andersen, Lene N.
; APPLICANT: Schulein, Martin
; APPLICANT: Lange, Niels Erik K.
; APPLICANT: Bjornvad, Mads E.
; APPLICANT: Schnorr, Kirk
; TITLE OF INVENTION: Pectin Degrading Enzymes From Bacillus
; FILE REFERENCE: 5377-200-US
; CURRENT APPLICATION NUMBER: US/09/670,141
; CURRENT FILING DATE: 2000-09-26
; PRIOR APPLICATION NUMBER: 09/198,956
; PRIOR FILING DATE: 1998-11-24
; PRIOR APPLICATION NUMBER: 1344/97
; PRIOR FILING DATE: 1997-11-24
; PRIOR APPLICATION NUMBER: 60/067,240
; PRIOR FILING DATE: 1997-12-02
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 7
; LENGTH: 1026
; TYPE: DNA
; ORGANISM: Bacillus licheniformis
US-09-670-141-7

Query Match 73.6%; Score 16.2; DB 4; Length 1026;
Best Local Similarity 85.7%; Pred. No. 36;
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 GACGTTGACACGGCGGTTTC 21
Db 51 GTCATTGACACGGCGGTTTC 71

RESULT 13

US-08-729-214-23
; Sequence 23, Application US/08729214
; Patent No. 5817502
; GENERAL INFORMATION:
; APPLICANT: Ligon, James M.
; APPLICANT: Hill, Dwight Steven
; APPLICANT: Ryals, John Andrew
; APPLICANT: Hammer, Phillip E.
; APPLICANT: van Pee, Karl-Heinz
; APPLICANT: Kirner, Sabine
; TITLE OF INVENTION: Genes for the synthesis of
; antipathogenic substances
; NUMBER OF SEQUENCES: 27
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Ciba-Geigy Corporation
; STREET: 520 White Plains Road

;; CITY: Tarrytown
;; STATE: NY
;; COUNTRY: USA
;; ZIP: 10591
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: PatentIn Release #1.0, Version #1.25
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/729,214
;; FILING DATE: TEA
;; CLASSIFICATION: 435
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Meigs, J. Timothy
;; REGISTRATION NUMBER: 38,241
;; REFERENCE/DOCKET NUMBER: CGC 1506/CIPS
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 919-541-8587
;; TELEFAX: 919-541-8689
;; INFORMATION FOR SEQ ID NO: 23:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 6250 base pairs
;; TYPE: nucleic acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: DNA (genomic)
;; HYPOTHETICAL: NO
;; FEATURE:
;; NAME/KEY: CDS
;; LOCATION: 615..2228
;; OTHER INFORMATION: /label= ORF1
;; OTHER INFORMATION: /note= "Open Reading Frame #1 of DNA sequence"
;; FEATURE:
;; NAME/KEY: CDS
;; LOCATION: 2231..3313
;; OTHER INFORMATION: /label= ORF2
;; OTHER INFORMATION: /note= "Open Reading Frame #2 of DNA sequence"
;; FEATURE:
;; NAME/KEY: CDS
;; LOCATION: 3368..5065
;; OTHER INFORMATION: /label= ORF3
;; OTHER INFORMATION: /note= "Open Reading Frame #3 of DNA sequence"
;; FEATURE:
;; NAME/KEY: CDS
;; LOCATION: 5093..6202
;; OTHER INFORMATION: /label= ORF4
;; OTHER INFORMATION: /note= "Open Reading Frame #4 of DNA sequence"
US-08-729-214-23

Query Match 73.6%; Score 16.2; DB 1; Length 6250;
Best Local Similarity 85.7%; Pred. No. 46;
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 GACGTTGACACGGCGGTTTC 21
Db 5555 GACGTCGACACGGCGGCTTC 5575

RESULT 14

US-09-028-934-23
; Sequence 23, Application US/09028934
; Patent No. 6117670
; GENERAL INFORMATION:
; APPLICANT: Ligon, James M.
; APPLICANT: Hill, Dwight S.
; APPLICANT: Lam, Steven T.
; APPLICANT: Hammer, Phillip E.
; APPLICANT: van Pee, Karl-Heinz
; APPLICANT: Kirner, Sabine
; APPLICANT: Young, Thomas R.
; TITLE OF INVENTION: Pyrrolnitrin Biosynthesis Genes and Uses
; TITLE OF INVENTION: Thereof

APPLICANT: Hammer, Philip E.
 APPLICANT: van Pee, Karl-Heinz
 APPLICANT: Kirner, Sabine
 APPLICANT: Young, Thomas R.
 TITLE OF INVENTION: Pyrolnitrin Biosynthesis Genes and Uses
 TITLE OF INVENTION: Pyrolnitrin Biosynthesis Genes and Uses
 NUMBER OF SEQUENCES: 37
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: NO. 6117670artis Corporation
 STREET: 3054 Cornwallis Road
 CITY: Research Triangle Park
 STATE: NC
 COUNTRY: USA
 ZIP: 27709
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/028,934
 FILING DATE:
 CLASSIFICATION:
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/729,214
 FILING DATE: 09-OCT-1996
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/258,261
 FILING DATE: 08-JUN-1994
 ATTORNEY/AGENT INFORMATION:
 NAME: Meigs, J. Timothy
 REGISTRATION NUMBER: 38,241
 REFERENCE/DOCKET NUMBER: CGC1506/CIP7
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 919-541-8587
 TELEFAX: 919-541-8689
 INFORMATION FOR SEQ ID NO: 28:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 8931 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: DNA (genomic)
 HYPOTHETICAL: NO
 ANTI-SENSE: NO
 ORIGINAL SOURCE:
 ORGANISM: Burkholdaria cepacia
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 657...2267
 OTHER INFORMATION: /product= "PrnA"
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 2270...3355
 OTHER INFORMATION: /product= "PrnB"
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 3421...5121
 OTHER INFORMATION: /product= "PrnC"
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 5145...6266
 OTHER INFORMATION: /product= "PrnD"
 US-09-028-934-28

QY 1 GACGTTGACAAACGGCGGTTTC 21
|||
Db 5555 GACGTCGACAAACGGCGACTTC 5575

RESULT 15
US-09-028-934-28
; Sequence 28, Application US/09028934
; Patent No. 6117670
; GENERAL INFORMATION:
; APPLICANT: Ligon, James M.
; APPLICANT: Hill, Dwight S.
; APPLICANT: Lam, Steven T.

Query Match 73.6%; Score 16.2; DB 3; Length 8931;
Best Local Similarity 85.7%; Pred. No. 48;
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 GACGTTGACAAACGGCGGTTC 21
|||||
Db 5613 GACGTCGACAAACGGCGACTTC 5633

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 18, 2003, 06:35:31 ; Search time 91.3333 Seconds
(without alignments)
241.975 Million cell updates/sec

Title: US-09-270-437D-14

Perfect score: 22

Sequence: 1 gacgttgacacggcggtttct 22

Scoring table: IDENTITY NUC

Gapop 10_0 , Gapext 1.0

Searched: 639749 seqs, 502280978 residues

Total number of hits satisfying chosen parameters: 1279498

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- Published Applications NA:*
- 1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
 - 2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
 - 3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
 - 4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
 - 5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
 - 6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
 - 7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
 - 8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
 - 9: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
 - 10: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
 - 11: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
 - 12: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
 - 13: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
 - 14: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. NO. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	22	100.0	3283	10	US-09-899-651-8
C 2	22	100.0	3412	10	US-09-899-651-6
C 3	22	100.0	3694	10	US-09-764-864-329
C 4	16.8	76.4	369	10	US-09-924-035A-3
C 5	16.8	76.4	486	9	US-09-938-842A-689
C 6	16.2	73.6	1036	12	US-10-072-152-3
C 7	16.2	73.6	1035	10	US-09-974-300-2650
C 8	16.2	73.6	1653	9	US-09-906-419-28
C 9	15.8	71.8	1035	9	US-10-169-953-1
C 10	15.8	71.8	1279	9	US-09-813-153-59
C 11	15.8	71.8	1462	9	US-09-764-852-516
C 12	15.8	71.8	1842	9	US-09-938-842A-1028
C 13	15.8	71.8	3057	10	US-09-974-300-2660
C 14	15.6	70.9	439	10	US-09-960-352-11807
C 15	15.6	70.9	431	10	US-09-867-701-568
C 16	15.6	70.9	531	9	US-09-738-626-1202
C 17	15.6	70.9	1276	9	US-09-982-704-12
C 18	15.6	70.9	1287	12	US-10-062-254-243
C 19	15.6	70.9	1335	9	US-09-738-626-3451

C 20	15.6	70.9	1338	9	US-10-067-974-11
C 21	15.6	70.9	1400	9	US-09-746-660A-43
C 22	15.6	70.9	1596	9	US-09-938-842A-785
C 23	15.6	70.9	3579	9	US-10-226-136-18
C 24	15.6	70.9	3309400	9	US-09-738-626-1
C 25	15.4	70.0	222	10	US-09-974-300-4137
C 26	15.2	69.1	266	10	US-09-878-574-7377
C 27	15.2	69.1	696	9	US-09-738-626-581
C 28	15.2	69.1	1176	10	US-09-974-300-1080
C 29	15.2	69.1	1330	10	US-09-974-300-2336
C 30	15.2	69.1	3798	10	US-09-974-300-2175
C 31	15.2	69.1	3798	10	US-09-974-300-2193
C 32	15.2	69.1	5134	9	US-09-902-432-1
C 33	15.2	69.1	6160	9	US-09-902-432-3
C 34	15.2	69.1	12860	10	US-09-070-927A-144
C 35	15.2	69.1	33023	10	US-09-880-107-7350
C 36	14.8	67.3	315	9	US-09-738-626-3285
C 37	14.8	67.3	555	10	US-09-974-300-5604
C 38	14.8	67.3	633	10	US-09-974-300-1122
C 39	14.8	67.3	944	10	US-09-770-445-351
C 40	14.8	67.3	1257	9	US-09-738-626-3137
C 41	14.8	67.3	1971	10	US-09-974-300-322
C 42	14.8	67.3	3044	9	US-09-927-827-32
C 43	14.8	67.3	111282	12	US-10-094-989-3
C 44	14.8	67.3	3309400	9	US-09-738-626-1
C 45	14.6	66.4	222	10	US-09-770-696-20

ALIGNMENTS

RESULT 1
US-09-899-651-8/c
; Sequence 8, Application US/09899651
; Patent No. US20020111470A1
; GENERAL INFORMATION:
; APPLICANT: Chen, Yao-Tseung
; APPLICANT: Gure, Ali
; APPLICANT: Tsang, Solam
; APPLICANT: Stockert, Elisabeth
; APPLICANT: Jager, Elke
; APPLICANT: Knuth, Alexander
; APPLICANT: Old, Lloyd J.
; TITLE OF INVENTION: Isolated Nucleic Acid Molecules Encoding Cancer
; TITLE OF INVENTION: Associated
; TITLE OF INVENTION: Antigen, The Antigens Per Se, And Uses Thereof
; FILE REFERENCE: LUD 5538
; CURRENT APPLICATION NUMBER: US/09/899,651
; PRIOR FILING DATE: 2001-07-06
; PRIOR APPLICATION NUMBER: US/09/061,709
; PRIOR FILING DATE: 1998-04-17
; NUMBER OF SEQ ID NOS: 8
; SEQ ID NO 8
; LENGTH: 3283
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
US-09-899-651-8

Query Match 100.0%; Score 22; DB 10; Length 3283;
Best Local Similarity 100.0%; Pred. No. 0.059;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GACGTTGACACGCGGTTTCT 22
Db 444 GACGTTGACACGCGGTTTCT 423

RESULT 2
US-09-899-651-6/c
; Sequence 6, Application US/09899651
; Patent No. US20020111470A1
; GENERAL INFORMATION:

; CURRENT APPLICATION NUMBER: US/10/072,152
 ; CURRENT FILING DATE: 2002-02-07
 ; PRIOR APPLICATION NUMBER: US/09/198,955
 ; PRIOR FILING DATE: 1998-11-24
 ; PRIOR APPLICATION NUMBER: 1343/97
 ; PRIOR FILING DATE: 1997-11-24
 ; PRIOR APPLICATION NUMBER: 1344/97
 ; PRIOR FILING DATE: 1997-11-24
 ; PRIOR APPLICATION NUMBER: 60/067,249
 ; PRIOR FILING DATE: 1997-12-02
 ; PRIOR APPLICATION NUMBER: 60/067,240
 ; PRIOR FILING DATE: 1997-12-02
 ; PRIOR APPLICATION NUMBER: 09/073,684
 ; PRIOR FILING DATE: 1998-05-06
 ; PRIOR APPLICATION NUMBER: 09/184,217
 ; PRIOR FILING DATE: 1998-11-02
 ; NUMBER OF SEQ ID NOS: 32
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 3
 ; LENGTH: 1026
 ; TYPE: DNA
 ; ORGANISM: Bacillus licheniformis - ATCC 14580
 US-10-072-152-3

Query Match 73.6%; Score 16.2; DB 12; Length 1026;
 Best Local Similarity 85.7%; Pred. No. 52;
 Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 Qy 1 GACGTTGACACGGCGGTTTC 21
 Db 51 GTCATTGACACGGCGGTTTC 71

RESULT 7
 US-09-974-300-2650
 ; Sequence 2650, Application US/09974300
 ; Patent No. US20020146721A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Berka, Randy M.
 ; TITLE OF INVENTION: Methods for Monitoring Multiple Gene
 ; FILE OF INVENTION: Expression
 ; FILE REFERENCE: 10085,500-US
 ; CURRENT APPLICATION NUMBER: US/09/974,300
 ; CURRENT FILING DATE: 2001-10-05
 ; PRIOR APPLICATION NUMBER: 09/680,598
 ; PRIOR FILING DATE: 2000-10-06
 ; PRIOR APPLICATION NUMBER: 60/279,526
 ; PRIOR FILING DATE: 2001-03-27
 ; NUMBER OF SEQ ID NOS: 8481
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 2650
 ; LENGTH: 1035
 ; TYPE: DNA
 ; ORGANISM: Bacillus licheniformis
 US-09-974-300-2650

Query Match 73.6%; Score 16.2; DB 10; Length 1035;
 Best Local Similarity 85.7%; Pred. No. 52;
 Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 Qy 1 GACGTTGACACGGCGGTTTC 21
 Db 62 GTCATTGACACGGCGGTTTC 82

RESULT 8
 US-09-906-419-28/c
 ; Sequence 28, Application US/09906419
 ; Publication No. US2003003757A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Shockey, Jay
 ; APPLICANT: Schnurr, Judy

; APPLICANT: Browse, John
 ; TITLE OF INVENTION: Plant Acyl-CoA Synthetases
 ; FILE REFERENCE: DOW-04679
 ; CURRENT APPLICATION NUMBER: US/09/906,419
 ; CURRENT FILING DATE: 2001-07-16
 ; PRIOR APPLICATION NUMBER: 60/220,474
 ; PRIOR FILING DATE: 2000-07-21
 ; NUMBER OF SEQ ID NOS: 120
 ; SOFTWARE: PatentIn version 3.0
 ; SEQ ID NO 28
 ; LENGTH: 1653
 ; TYPE: DNA
 ; ORGANISM: Arabidopsis thaliana
 US-09-906-419-28

Query Match 73.6%; Score 16.2; DB 9; Length 1653;
 Best Local Similarity 85.7%; Pred. No. 56;
 Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 Qy 1 GACGTTGACACGGCGGTTTC 21
 Db 1014 GACGATTACACGGCGGTTTC 994

RESULT 9
 US-10-169-953-1
 ; Sequence 1, Application US/10169953
 ; Publication No. US20030044915A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Thierry BAUSSANT
 ; APPLICANT: Pascale JEANNIN
 ; APPLICANT: Yves DELNESTE
 ; APPLICANT: Françoise LAMNY
 ; APPLICANT: Jean-Yves BONNEFOY
 ; TITLE OF INVENTION: METHOD FOR PREPARING A POLYPEPTIDE SOLUBLE IN AN AQUEOUS SOLVENT
 ; FILE OF INVENTION: IN THE ABSENCE OF DETERGENT
 ; FILE REFERENCE: D18390
 ; CURRENT APPLICATION NUMBER: US/10/169,953
 ; CURRENT FILING DATE: 2002-07-03
 ; PRIOR APPLICATION NUMBER: FR 00 00070
 ; PRIOR FILING DATE: 2000-01-04
 ; NUMBER OF SEQ ID NOS: 2
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 1
 ; LENGTH: 1035
 ; TYPE: DNA
 ; ORGANISM: Klebsiella pneumoniae
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (1)..(1032)
 ; FEATURE:
 ; OTHER INFORMATION: p40
 US-10-169-953-1

Query Match 71.8%; Score 15.8; DB 9; Length 1035;
 Best Local Similarity 89.5%; Pred. No. 84;
 Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 Qy 3 CGTTGACACGGCGGTTTC 21
 Db 240 CGTTGACACGGCGGTTTC 258

RESULT 10
 US-09-813-153-59
 ; Sequence 59, Application US/09813153
 ; Publication No. US20030045459A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Rosen et al.
 ; TITLE OF INVENTION: 67 Human secreted proteins
 ; FILE REFERENCE: P2023
 ; CURRENT APPLICATION NUMBER: US/09/813,153
 ; CURRENT FILING DATE: 2001-03-21

; GENERAL INFORMATION:
 ; APPLICANT: Warren, Wesley C.
 ; APPLICANT: Tao, Nengbing
 ; APPLICANT: Byatt, John C.
 ; APPLICANT: Mathialagan, Nagappan
 ; TITLE OF INVENTION: NUCLEIC ACID
 ; TITLE OF INVENTION: MUSCLE AND F
 ; FILE REFERENCE: 16511.006/37-21 (11
 ; CURRENT APPLICATION NUMBER: US/09
 ; CURRENT FILING DATE: 2001-09-24

```

RESULT 12
US-09-938-842A-1028
; Sequence 1028, Application US/09938842A
; Patent No. US20020160378A1
; GENERAL INFORMATION:
; APPLICANT: Harper, Jeff
; APPLICANT: Krebs, Joel
; APPLICANT: Wang, Xun
; APPLICANT: Zhu, Tong
; TITLE OF INVENTION: STRESS-REGULATED G

```

; NUMBER OF SEQ ID NOS: 15112
 ; SEQ ID NO 11807
 ; LENGTH: 411
 ; TYPE: DNA
 ; ORGANISM: Bos taurus
 ; OTHER INFORMATION: Clone ID: 50-LIB34-083-Q1-E1-E6
 US-09-960-352-11807

Query Match 70.9%; Score 15.6; DB 10; Length 411;
 Best Local Similarity 81.8%; Pred. No. 95;
 Matches 18; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

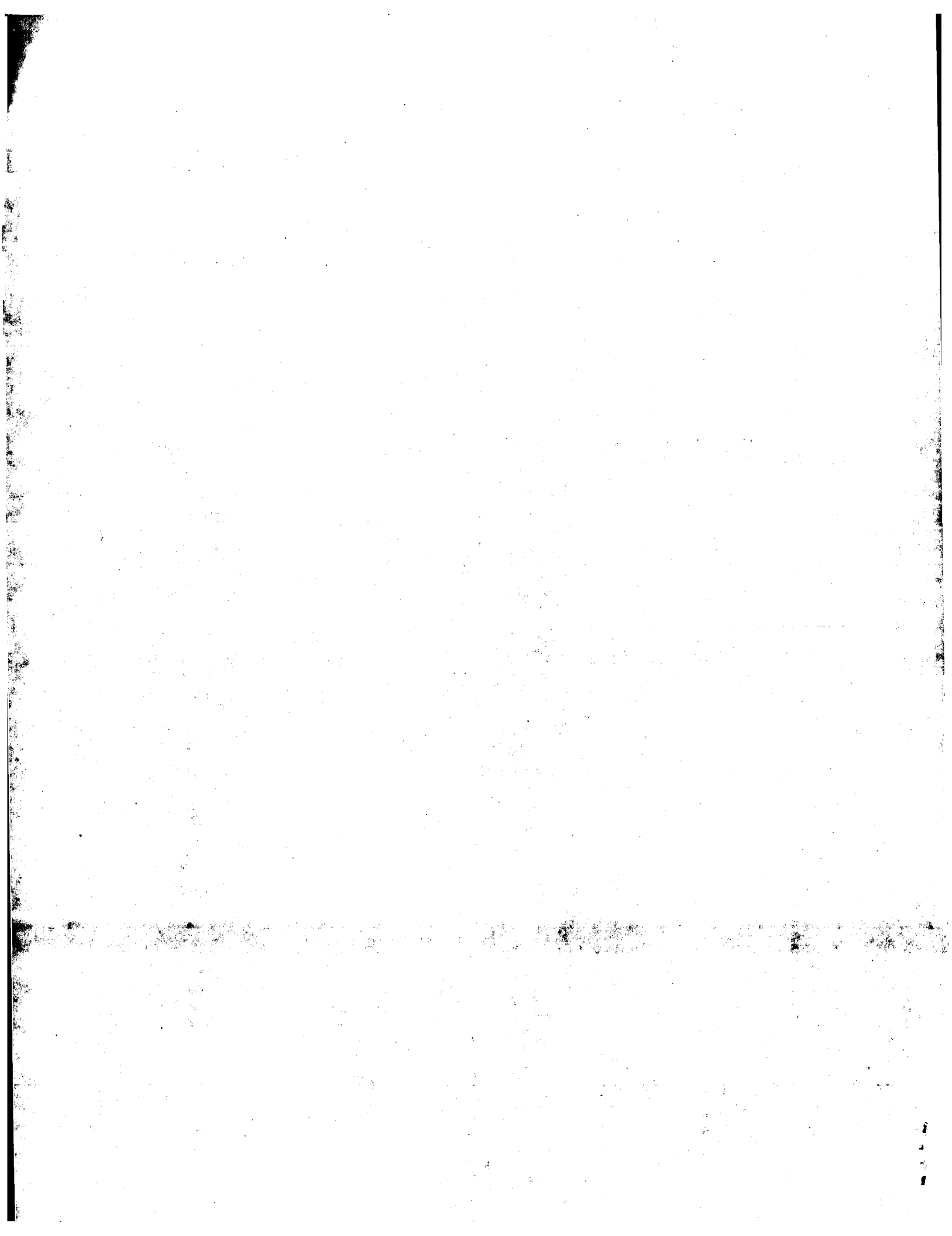
Qy 1 GACGTTGACACGCGGTTTCT 22
 Db 184 GACGTTGACGACTGCTGTTCT 205

RESULT 15
 US-09-867-701-568/c
 ; Sequence 568, Application US/09867701
 ; Patent No. US20020132237A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Aglate, Paul A.
 ; APPLICANT: Jones, Robert
 ; APPLICANT: Harlocker, Susan L.
 ; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
 ; TITLE OF INVENTION: AND DIAGNOSIS OF OVARIAN CANCER
 ; FILE REFERENCE: 210121.497
 ; CURRENT APPLICATION NUMBER: US/09/867,701
 ; CURRENT FILING DATE: 2001-05-29
 ; NUMBER OF SEQ ID NOS: 10912
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 568
 ; LENGTH: 439
 ; TYPE: DNA
 ; ORGANISM: Homo sapien
 ; FEATURE:
 ; NAME/KEY: misc.feature
 ; LOCATION: (1)...(439)
 ; OTHER INFORMATION: n = A,T,C or G
 US-09-867-701-568

Query Match 70.9%; Score 15.6; DB 10; Length 439;
 Best Local Similarity 81.8%; Pred. No. 95;
 Matches 18; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 GACGTTGACACGCGGTTTCT 22
 Db 26 GACATTGAGACACGAGTTTCT 5

Search completed: April 18, 2003, 10:16:07
 Job time : 96.5 secs



Mon Apr 21 10:30:38 2003

us-09-270-437d-14.rni

Page 7

Search completed: April 18, 2003, 07:34:42
Job time : 30.5 secs

GenCore version 5.1.1.4 p5 4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 18, 2003, 05:48:17 ; Search time 1211 Seconds
(without alignments)
456.759 Million cell updates/sec

Title: US-09-270-437D-14

Perfect score: 22

Sequence: 1 gacgtgacaacggcggtttct 22

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 24791104 seqs, 12571243825 residues

Total number of hits satisfying chosen parameters: 49582208

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: Pending Patents NA Main:*

2: /cgn2_6/ptodata/1/pna/PCTUS COMB.seq.*

3: /cgn2_6/ptodata/1/pna/US06 COMB.seq.*

4: /cgn2_6/ptodata/1/pna/US07 COMB.seq.*

5: /cgn2_6/ptodata/1/pna/US08 COMB.seq.*

6: /cgn2_6/ptodata/1/pna/US081 COMB.seq.*

7: /cgn2_6/ptodata/1/pna/US082 COMB.seq.*

8: /cgn2_6/ptodata/1/pna/US083 COMB.seq.*

9: /cgn2_6/ptodata/1/pna/US084 COMB.seq.*

10: /cgn2_6/ptodata/1/pna/US085 COMB.seq.*

11: /cgn2_6/ptodata/1/pna/US086 COMB.seq.*

12: /cgn2_6/ptodata/1/pna/US087 COMB.seq.*

13: /cgn2_6/ptodata/1/pna/US088 COMB.seq.*

14: /cgn2_6/ptodata/1/pna/US089 COMB.seq.*

15: /cgn2_6/ptodata/1/pna/US090 COMB.seq.*

16: /cgn2_6/ptodata/1/pna/US091 COMB.seq.*

17: /cgn2_6/ptodata/1/pna/US092 COMB.seq.*

18: /cgn2_6/ptodata/1/pna/US093 COMB.seq.*

19: /cgn2_6/ptodata/1/pna/US094 COMB.seq.*

20: /cgn2_6/ptodata/1/pna/US095A COMB.seq.*

21: /cgn2_6/ptodata/1/pna/US095B COMB.seq.*

22: /cgn2_6/ptodata/1/pna/US095C COMB.seq.*

23: /cgn2_6/ptodata/1/pna/US095D COMB.seq.*

24: /cgn2_6/ptodata/1/pna/US096A COMB.seq.*

25: /cgn2_6/ptodata/1/pna/US096B COMB.seq.*

26: /cgn2_6/ptodata/1/pna/US096C COMB.seq.*

27: /cgn2_6/ptodata/1/pna/US096D COMB.seq.*

28: /cgn2_6/ptodata/1/pna/US096E COMB.seq.*

29: /cgn2_6/ptodata/1/pna/US097A COMB.seq.*

30: /cgn2_6/ptodata/1/pna/US097B COMB.seq.*

31: /cgn2_6/ptodata/1/pna/US097C COMB.seq.*

32: /cgn2_6/ptodata/1/pna/US098A COMB.seq.*

33: /cgn2_6/ptodata/1/pna/US098B COMB.seq.*

34: /cgn2_6/ptodata/1/pna/US098C COMB.seq.*

35: /cgn2_6/ptodata/1/pna/US099A COMB.seq.*

36: /cgn2_6/ptodata/1/pna/US099B COMB.seq.*

37: /cgn2_6/ptodata/1/pna/US099C COMB.seq.*

38: /cgn2_6/ptodata/1/pna/US099D COMB.seq.*

39: /cgn2_6/ptodata/1/pna/US100A COMB.seq.*

40: /cgn2_6/ptodata/1/pna/US100B COMB.seq.*

41: /cgn2_6/ptodata/1/pna/US101A COMB.seq.*

42: /cgn2_6/ptodata/1/pna/US101B COMB.seq.*

43: /cgn2_6/ptodata/1/pna/US102B COMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	22	100.0	272	13	US-08-359-395-2440
C 2	22	100.0	272	20	US-09-539-334-17512
C 3	22	100.0	347	25	US-09-652-121-1684
C 4	22	100.0	349	16	US-09-289-768-12874
C 5	22	100.0	349	35	US-09-939-397-12874
C 6	22	100.0	363	16	US-09-289-768-32161
C 7	22	100.0	363	35	US-09-939-397-32161
C 8	22	100.0	409	25	US-09-652-918-3592
C 9	22	100.0	465	17	US-09-304-649-680
C 10	22	100.0	465	32	US-09-840-424-680
C 11	22	100.0	467	19	US-09-528-409-49810
C 12	22	100.0	467	35	US-09-933-524-49810
C 13	22	100.0	467	35	US-09-933-524-49810
C 14	22	100.0	546	56	US-60-125-787-430
C 15	22	100.0	2010	1	PCT-US01-08631-6785
C 16	22	100.0	3283	16	US-09-270-437-8
C 17	22	100.0	3283	33	US-09-899-651-8
C 18	22	100.0	3374	1	PCT-US01-01307-113
C 19	22	100.0	3374	39	US-10-092-302-113
C 20	22	100.0	3412	16	US-09-270-437-6
C 21	22	100.0	3412	33	US-09-899-651-6
C 22	22	100.0	3412	33	US-09-899-651-6

44: /cgn2_6/ptodata/1/pna/US6000 COMB.seq.*

45: /cgn2_6/ptodata/1/pna/US6001 COMB.seq.*

46: /cgn2_6/ptodata/1/pna/US6002 COMB.seq.*

47: /cgn2_6/ptodata/1/pna/US6003 COMB.seq.*

48: /cgn2_6/ptodata/1/pna/US6004 COMB.seq.*

49: /cgn2_6/ptodata/1/pna/US6005 COMB.seq.*

50: /cgn2_6/ptodata/1/pna/US6006 COMB.seq.*

51: /cgn2_6/ptodata/1/pna/US6007 COMB.seq.*

52: /cgn2_6/ptodata/1/pna/US6008 COMB.seq.*

53: /cgn2_6/ptodata/1/pna/US6009 COMB.seq.*

54: /cgn2_6/ptodata/1/pna/US6010 COMB.seq.*

55: /cgn2_6/ptodata/1/pna/US6011 COMB.seq.*

56: /cgn2_6/ptodata/1/pna/US6012 COMB.seq.*

57: /cgn2_6/ptodata/1/pna/US6013 COMB.seq.*

58: /cgn2_6/ptodata/1/pna/US6014 COMB.seq.*

59: /cgn2_6/ptodata/1/pna/US6015 COMB.seq.*

60: /cgn2_6/ptodata/1/pna/US6016 COMB.seq.*

61: /cgn2_6/ptodata/1/pna/US6017 COMB.seq.*

62: /cgn2_6/ptodata/1/pna/US6018 COMB.seq.*

63: /cgn2_6/ptodata/1/pna/US6019 COMB.seq.*

64: /cgn2_6/ptodata/1/pna/US6020 COMB.seq.*

65: /cgn2_6/ptodata/1/pna/US6021 COMB.seq.*

66: /cgn2_6/ptodata/1/pna/US6022 COMB.seq.*

67: /cgn2_6/ptodata/1/pna/US6023 COMB.seq.*

68: /cgn2_6/ptodata/1/pna/US6024 COMB.seq.*

69: /cgn2_6/ptodata/1/pna/US6025 COMB.seq.*

70: /cgn2_6/ptodata/1/pna/US6026 COMB.seq.*

71: /cgn2_6/ptodata/1/pna/US6027 COMB.seq.*

72: /cgn2_6/ptodata/1/pna/US6028 COMB.seq.*

73: /cgn2_6/ptodata/1/pna/US6029 COMB.seq.*

74: /cgn2_6/ptodata/1/pna/US6030 COMB.seq.*

75: /cgn2_6/ptodata/1/pna/US6031 COMB.seq.*

76: /cgn2_6/ptodata/1/pna/US6032 COMB.seq.*

77: /cgn2_6/ptodata/1/pna/US6033 COMB.seq.*

78: /cgn2_6/ptodata/1/pna/US6034 COMB.seq.*

79: /cgn2_6/ptodata/1/pna/US6035 COMB.seq.*

80: /cgn2_6/ptodata/1/pna/US6036 COMB.seq.*

81: /cgn2_6/ptodata/1/pna/US6037 COMB.seq.*

82: /cgn2_6/ptodata/1/pna/US6038 COMB.seq.*

83: /cgn2_6/ptodata/1/pna/US6039 COMB.seq.*

84: /cgn2_6/ptodata/1/pna/US6040 COMB.seq.*

85: /cgn2_6/ptodata/1/pna/US6041 COMB.seq.*

Sequence 27652, A
Sequence 6786, Ap
Sequence 146, App
Sequence 146, App
Sequence 48, App1
Sequence 329, App
Sequence 329, App
Sequence 329, App
Sequence 8906, Ap
Sequence 9743, Ap
Sequence 6556, Ap
Sequence 20241, A
Sequence 20241, A
Sequence 3571, A
Sequence 45804, A
Sequence 514, App
Sequence 377, App
Sequence 1680, Ap
Sequence 133, App
Sequence 44, App1
Sequence 45836, A
Sequence 11032, A
Sequence 76081, A
Sequence 60934, A

US-60-172-360-27652
1 PCT-US01-08631-6786
1 PCT-US02-07826-146
39 US-10-097-340-146
84 US-60-408-385-48
1 PCT-US01-01341-329
30 US-09-764-864-329
39 US-10-080-129-329
25 US-09-644-869-8906
27 US-09-699-998-9743
22 US-09-721-589-6556
29 US-09-814-353-20241
31 US-09-814-353A-20241
35 US-09-652-355-3571
21 95.5 492 25 US-09-540-212A-45804
29 92.7 289 21 US-09-677-417-514
30 92.7 350 26 US-09-535-356-377
20 92.7 455 20 US-09-535-356-377
38 20.4 92.7 386 26 US-09-677-417-1680
39 19.4 88.2 386 26 US-09-460-595-133
40 19 86.4 462 18 US-09-188-162-44
41 18.8 85.5 698 62 US-09-815-264-45836
42 18.8 85.5 1501 31 US-09-702-134-11032
43 18.8 85.5 36198 28 US-09-815-264-76081
44 18.8 85.5 36198 31 US-09-815-264-76081
45 18.8 85.5 36241 24 US-09-620-392-60934

ALIGNMENTS

RESULT 1
US-08-959-395-2440/c
; Sequence 2440, Application US/08959395
; GENERAL INFORMATION:
; APPLICANT: Gooding, Douglas H.
; APPLICANT: Stuve, Laura L.
; APPLICANT: Stuart, Susan G.
; APPLICANT: Ito, Laura Y.
; APPLICANT: Aketblom, Ingrid B.
; APPLICANT: Delegeane, Angelo M.
; APPLICANT: Naughton, Rebecca E.
; APPLICANT: Klingler, Tod M.
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM
; TITLE OF INVENTION: HUMAN PANCREATIC ISLET CELLS
; NUMBER OF SEQUENCES: 5789
; CORRESPONDENCE ADDRESS:
; ADDRESS: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/959,395
; FILING DATE: HEREWITH
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/030,755
; FILING DATE: OCTOBER 28, 1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/033,551
; FILING DATE: DECEMBER 20, 1996
; ATTORNEY/AGENT INFORMATION:
; NAME: CERONE, MICHAEL C.
; REGISTRATION NUMBER: 39,132
; REFERENCE/DOCKET NUMBER: PD-0259 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 855-0555
; TELEFAX: (650) 845-4166

INFORMATION FOR SEQ ID NO: 2440:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 272 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; IMMEDIATE SOURCE:
; CLONE: 2077095H2
US-08-959-395-2440
Query Match 100.0%; Score 22; DB 13; Length 272;
Best Local Similarity 100.0%; Pred. No. 1.6; Indels 0; Gaps 0;
Matches 22; Conservative 0; Mismatches 0;
QY 1 GACGTTGACACGCGGTTTCT 22
DB 179 GACGTTGACACGCGGTTTCT 158
RESULT 2
US-09-539-334-17512/c
; Sequence 17512, Application US/09539334
; GENERAL INFORMATION:
; APPLICANT: Seilhamer, Jeffrey J.
; APPLICANT: Delegeane, Angelo M.
; APPLICANT: Stuart, Susan G.
; APPLICANT: Stuve, Laura L.
; APPLICANT: Mullahy, Sara J.
; APPLICANT: Naughton, Rebecca E.
; TITLE OF INVENTION: POLYNUCLEOTIDES OF ENDOCRINE SYSTEM
; FILE REFERENCE: PD-1026 CIP
; CURRENT APPLICATION NUMBER: US/09/539,334
; CURRENT FILING DATE: 2000-03-30
; "Prior application data removed - refer to PALM or file wrapper"
; NUMBER OF SEQ ID NOS: 38381
; SOFTWARE: PERL Program
; SEQ ID NO 17512
; LENGTH: 272
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No: hu00470097
US-09-539-334-17512
Query Match 100.0%; Score 22; DB 20; Length 272;
Best Local Similarity 100.0%; Pred. No. 1.6; Indels 0; Gaps 0;
Matches 22; Conservative 0; Mismatches 0;
QY 1 GACGTTGACACGCGGTTTCT 22
DB 179 GACGTTGACACGCGGTTTCT 158
RESULT 3
US-09-652-121-1684/c
; Sequence 1684, Application US/09652121
; GENERAL INFORMATION:
; APPLICANT: Distefano, Peter
; TITLE OF INVENTION: NOVEL NUCLEIC ACID MOLECULES AND USES
; FILE REFERENCE: 1600.1188-001
; CURRENT APPLICATION NUMBER: US/09/652,121
; CURRENT FILING DATE: 2000-08-30
; PRIOR APPLICATION NUMBER: 60/151,129
; PRIOR FILING DATE: 1999-08-30
; NUMBER OF SEQ ID NOS: 7615
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1684
; LENGTH: 347
; TYPE: DNA
; ORGANISM: Homo sapiens

FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)...(347)
OTHER INFORMATION: n = A,T,C or G
US-09-652-121-1684

Query Match 100.0%; Score 22; DB 25; Length 347;
Best Local Similarity 100.0%; Pred. No. 1.7;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACGTTGACAAACGGCGGTTTCT 22
|||||
DB 323 GACGTTGACAAACGGCGGTTTCT 302

RESULT 4

US-09-289-768-12874/c
Sequence 12874, Application US/09289768
GENERAL INFORMATION:
APPLICANT: Hyseq, Inc.
TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
FILE REFERENCE: 20411-765
CURRENT APPLICATION NUMBER: US/09/289,768
CURRENT FILING DATE: 1999-04-08
NUMBER OF SEQ ID NOS: 39996
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 12874
LENGTH: 349
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)...(349)
OTHER INFORMATION: n = A,T,C or G
US-09-289-768-12874

Query Match 100.0%; Score 22; DB 16; Length 349;
Best Local Similarity 100.0%; Pred. No. 1.7;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACGTTGACAAACGGCGGTTTCT 22
|||||
DB 122 GACGTTGACAAACGGCGGTTTCT 101

RESULT 5

US-09-939-397-12874/c
Sequence 12874, Application US/09939397
GENERAL INFORMATION:
APPLICANT: Hyseq, Inc.
TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
FILE REFERENCE: 20411-765
CURRENT APPLICATION NUMBER: US/09/939,397
CURRENT FILING DATE: 2001-08-24
PRIOR FILING DATE: 1999-04-08
NUMBER OF SEQ ID NOS: 39996
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 12874
LENGTH: 349
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)...(349)
OTHER INFORMATION: n = A,T,C or G
US-09-939-397-12874

Query Match 100.0%; Score 22; DB 35; Length 349;
Best Local Similarity 100.0%; Pred. No. 1.7;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACGTTGACAAACGGCGGTTTCT 22
|||||
DB 122 GACGTTGACAAACGGCGGTTTCT 101

RESULT 6

US-09-289-768-32161/c
Sequence 32161, Application US/09289768
GENERAL INFORMATION:
APPLICANT: Hyseq, Inc.
TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
FILE REFERENCE: 20411-765
CURRENT APPLICATION NUMBER: US/09/289,768
CURRENT FILING DATE: 1999-04-08
NUMBER OF SEQ ID NOS: 39996
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 32161
LENGTH: 363
TYPE: DNA
ORGANISM: Homo sapiens
US-09-289-768-32161

Query Match 100.0%; Score 22; DB 16; Length 363;
Best Local Similarity 100.0%; Pred. No. 1.7;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACGTTGACAAACGGCGGTTTCT 22
|||||
DB 276 GACGTTGACAAACGGCGGTTTCT 255

RESULT 7

US-09-939-397-32161/c
Sequence 32161, Application US/09939397
GENERAL INFORMATION:
APPLICANT: Hyseq, Inc.
TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
FILE REFERENCE: 20411-765
CURRENT APPLICATION NUMBER: US/09/939,397
CURRENT FILING DATE: 2001-08-24
PRIOR FILING DATE: 1999-04-08
NUMBER OF SEQ ID NOS: 39996
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 32161
LENGTH: 363
TYPE: DNA
ORGANISM: Homo sapiens
US-09-939-397-32161

Query Match 100.0%; Score 22; DB 35; Length 363;
Best Local Similarity 100.0%; Pred. No. 1.7;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACGTTGACAAACGGCGGTTTCT 22
|||||
DB 276 GACGTTGACAAACGGCGGTTTCT 255

RESULT 8

US-09-652-918-3592/c
Sequence 3592, Application US/09652918
GENERAL INFORMATION:
APPLICANT: Galvin, Katherine
APPLICANT: Holtzman, Douglas A.
TITLE OF INVENTION: NOVEL NUCLEIC ACID MOLECULES AND USES
FILE REFERENCE: 1600.1187-001
CURRENT APPLICATION NUMBER: US/09/652,918
CURRENT FILING DATE: 2000-08-30

Mon Apr 21 10:30:43 2003

```

; PRIOR FILING DATE: 1998-05-04
; PRIOR APPLICATION NUMBER: US 60/123,523
; PRIOR FILING DATE: 1999-03-08
; PRIOR APPLICATION NUMBER: US 60/126,974
; PRIOR FILING DATE: 1999-03-30
; PRIOR APPLICATION NUMBER: US 09/304,649
; PRIOR FILING DATE: 1999-05-04
; NUMBER OF SEQ ID NOS: 1311
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 680
; LENGTH: 465
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (1)...(465)
; OTHER INFORMATION: n = A,T,C or G
US-09-840-424-680

Query Match      100.0%; Score 22; DB 32; Length 465;
Best Local Similarity 100.0%; Pred. No. 1.7;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACGTTGACACGCGGTTTCT 22
    |||||
Db 464 GACGTTGACACGCGGTTTCT 443

RESULT 11
US-09-528-409-49810/c
; Sequence 49810, Application US/09528409
; GENERAL INFORMATION:
; APPLICANT: Drmanac, Radoje T.
; APPLICANT: Labat, Ivan
; APPLICANT: Stache-Crain, Birgit
; APPLICANT: Dickson, Mark
; APPLICANT: Jones, Lee W.
; TITLE OF INVENTION: Novel Nucleic Acid Sequences Obtained
; FILE REFERENCE: 774
; CURRENT APPLICATION NUMBER: US/09/528,409
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: 60/125,453
; PRIOR FILING DATE: 1999-03-19
; NUMBER OF SEQ ID NOS: 116231
; SOFTWARE: HY-patent.pl Version 3.1
; SEQ ID NO 49810
; LENGTH: 467
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (1)...(467)
; OTHER INFORMATION: n = A,T,C or G
US-09-528-409-49810

Query Match      100.0%; Score 22; DB 19; Length 467;
Best Local Similarity 100.0%; Pred. No. 1.7;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACGTTGACACGCGGTTTCT 22
    |||||
Db 457 GACGTTGACACGCGGTTTCT 436

RESULT 12
US-09-933-524-49810/c
; Sequence 49810, Application US/09933524
; GENERAL INFORMATION:
; APPLICANT: Drmanac, Radoje T.
; APPLICANT: Labat, Ivan
; APPLICANT: Stache-Crain, Birgit
; APPLICANT: Dickson, Mark
; PRIOR APPLICATION NUMBER: US 60/084,098
; CURRENT FILING DATE: 2001-04-23
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3592
; LENGTH: 409
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (1)...(409)
; OTHER INFORMATION: n = A,T,C or G
US-09-652-918-3592

Query Match      100.0%; Score 22; DB 25; Length 409;
Best Local Similarity 100.0%; Pred. No. 1.7;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACGTTGACACGCGGTTTCT 22
    |||||
Db 262 GACGTTGACACGCGGTTTCT 241

RESULT 9
US-09-304-649-680/c
; Sequence 680, Application US/09304649
; GENERAL INFORMATION:
; APPLICANT: Gearing, David P.
; APPLICANT: McCarthy, Sean A.
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES DERIVED FROM A
; FILE REFERENCE: MLN98-10PM
; CURRENT APPLICATION NUMBER: US/09/304,649
; CURRENT FILING DATE: 1999-05-04
; EARLIER APPLICATION NUMBER: 60/084,098
; EARLIER FILING DATE: 1998-05-04
; EARLIER APPLICATION NUMBER: 60/123,523
; EARLIER FILING DATE: 1999-03-08
; EARLIER APPLICATION NUMBER: 60/126,974
; EARLIER FILING DATE: 1999-03-30
; NUMBER OF SEQ ID NOS: 1311
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 680
; LENGTH: 465
; TYPE: DNA
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (1)...(465)
; OTHER INFORMATION: n = A,T,C or G
US-09-304-649-680

Query Match      100.0%; Score 22; DB 17; Length 465;
Best Local Similarity 100.0%; Pred. No. 1.7;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACGTTGACACGCGGTTTCT 22
    |||||
Db 464 GACGTTGACACGCGGTTTCT 443

RESULT 10
US-09-840-424-680/c
; Sequence 680, Application US/09840424
; GENERAL INFORMATION:
; APPLICANT: Gearing, David P.
; APPLICANT: McCarthy, Sean A.
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES DERIVED FROM A
; FILE REFERENCE: 1600.1010-002
; CURRENT APPLICATION NUMBER: US/09/840,424
; CURRENT FILING DATE: 2001-04-23
; PRIOR APPLICATION NUMBER: US 60/084,098
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3592
; LENGTH: 409
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (1)...(409)
; OTHER INFORMATION: n = A,T,C or G
US-09-652-918-3592
```

; APPLICANT: Jones, Lee W.
; TITLE OF INVENTION: Novel Nucleic Acid Sequences Obtained
; FILE REFERENCE: 774
; CURRENT APPLICATION NUMBER: US/09/933,524
; PRIOR FILING DATE: 2001-08-20
; PRIOR APPLICATION NUMBER: 09/528,409
; PRIOR FILING DATE: 2000-03-17
; NUMBER OF SEQ ID NOS: 116231
; SOFTWARE: Hy-patent.pl Version 3.1
; SEQ ID NO 49810
; LENGTH: 467
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(467)
; OTHER INFORMATION: n = A,T,C or G
US-09-933-524-49810

Query Match 100.0%; Score 22; DB 35; Length 467;
Best Local Similarity 100.0%; Pred. No. 1.7;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACGTTGACACGCGGTTTCT 22
|||||
Db 457 GACGTTGACACGCGGTTTCT 436

RESULT 13
US-09-933-524A-49810/c
; Sequence 49810, Application US/09933524A
; GENERAL INFORMATION:
; APPLICANT: Drmanac, Radoje T.
; APPLICANT: Labat, Ivan
; APPLICANT: Stache-Crain, Birgit
; APPLICANT: Dickson, Mark
; APPLICANT: Jones, Lee W.
; TITLE OF INVENTION: Novel Nucleic Acid Sequences Obtained
; FILE REFERENCE: 774
; CURRENT APPLICATION NUMBER: US/09/933,524A
; PRIOR FILING DATE: 2001-08-20
; PRIOR APPLICATION NUMBER: 09/528,409
; PRIOR FILING DATE: 2000-05-17
; NUMBER OF SEQ ID NOS: 116231
; SOFTWARE: Hy-patent.pl Version 3.1
; SEQ ID NO 49810
; LENGTH: 467
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(467)
; OTHER INFORMATION: n = A,T,C or G
US-09-933-524A-49810

Query Match 100.0%; Score 22; DB 35; Length 467;
Best Local Similarity 100.0%; Pred. No. 1.7;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACGTTGACACGCGGTTTCT 22
|||||
Db 457 GACGTTGACACGCGGTTTCT 436

RESULT 14
US-60-125-787-430/c
; Sequence 430, Application US/60125787
; GENERAL INFORMATION:
; APPLICANT: Hodgson, David M.
; APPLICANT: Lincoln, Stephen E.
; APPLICANT: Russo, Frank D.

; APPLICANT: Spiro, Peter A.
; APPLICANT: Banville, Steve C.
; APPLICANT: Bratcher, Shawn R.
; APPLICANT: Dufour, Gerard E.
; APPLICANT: Cohen, Howard J.
; APPLICANT: Rosen, Bruce
; APPLICANT: Shah, Purvi
; APPLICANT: Chalup, Michael S.
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Jones, Anissa L.
; APPLICANT: Yu, Jimmy Y.
; APPLICANT: Greenawalt, Lila B.
; APPLICANT: Panzer, Scott R.
; APPLICANT: Roseberry, Ana M.
; APPLICANT: Wright, Rachel J.
; TITLE OF INVENTION: TRANSCRIPTION FACTOR MOLECULES
; FILE REFERENCE: PT-0017 P
; CURRENT APPLICATION NUMBER: US/60/125,787
; CURRENT FILING DATE: 1999-03-23
; NUMBER OF SEQ ID NOS: 591
; SOFTWARE: PERL Program
; SEQ ID NO 430
; LENGTH: 546
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: 22091.1
US-60-125-787-430

Query Match 100.0%; Score 22; DB 56; Length 546;
Best Local Similarity 100.0%; Pred. No. 1.7;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACGTTGACACGCGGTTTCT 22
|||||
Db 453 GACGTTGACACGCGGTTTCT 432

RESULT 15
PCT-US01-08631-6785/c
; Sequence 6785, Application PC/TUS0108631
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 21272-049
; CURRENT APPLICATION NUMBER: PCT/US01/08631
; CURRENT FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: 09/540,217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 09/649,167
; PRIOR FILING DATE: 2000-08-23
; NUMBER OF SEQ ID NOS: 60736
; SOFTWARE: Custom
; SEQ ID NO 6785
; LENGTH: 2010
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIMILAR
; LOCATION: (50)...(1843)
; OTHER INFORMATION: 100% homologous to Homo sapiens IGF-II mRNA-binding protein
; OTHER INFORMATION: 2,accession number AF117107,Smith-Waterman Scores=3059.
PCT-US01-08631-6785

Query Match 100.0%; Score 22; DB 1; Length 2010;
Best Local Similarity 100.0%; Pred. No. 1.9;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACGTTGACACGCGGTTTCT 22
|||||
Db 421 GACGTTGACACGCGGTTTCT 400

Mon Apr 21 10:30:43 2003

us-09-270-437d-14.rnp

Page 6

Search completed: April 18, 2003, 09:35:49
Job time : 1213 secs

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 18, 2003, 06:35:28 ; Search time 245.167 Seconds
(without alignments)
400.770 Million cell updates/sec

Title: US-09-270-437D-14

Perfect score: 22

Sequence: 1 gacgttgacacggcggtttct 22

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 5897297 seqs, 2233080881 residues

Total number of hits satisfying chosen parameters: 11794594

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Pending Patents NA New:

- 1: /cgn2_6/ptodata/1/pna/PCT_NEW_COMB.seq.*
- 2: /cgn2_6/ptodata/1/pna/US06_NEW_COMB.seq.*
- 3: /cgn2_6/ptodata/1/pna/US07_NEW_COMB.seq.*
- 4: /cgn2_6/ptodata/1/pna/US08_NEW_COMB.seq.*
- 5: /cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq.*
- 6: /cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq2.*
- 7: /cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq3.*
- 8: /cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq.*
- 9: /cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq2.*
- 10: /cgn2_6/ptodata/1/pna/US60_NEW_COMB.seq.*
- 11: /cgn2_6/ptodata/1/pna/US60_NEW_COMB.seq2.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	22	100.0	22	5	US-09-270-437D-14
2	22	100.0	3283	5	US-09-270-437D-8
3	22	100.0	3412	5	US-09-270-437D-6
4	22	100.0	3427	8	US-10-170-235-21870
5	22	100.0	3427	11	US-60-453-135-3834
6	22	100.0	3427	11	US-60-453-135-3834
7	22	100.0	3475	8	US-10-170-235-22086
8	22	100.0	3475	11	US-60-453-135-3835
9	22	100.0	3475	11	US-60-453-135-3835
10	18.8	85.5	201	11	US-60-452-680-30351
11	18.8	85.5	201	11	US-60-452-680-30361
12	18.8	85.5	201	11	US-60-453-135-20375
13	18.8	85.5	201	11	US-60-453-135-20375
14	18.8	85.5	201	11	US-60-453-135-20375
15	18.8	85.5	201	11	US-60-453-135-20375
16	18.8	85.5	201	11	US-60-453-135-20375
17	18.8	85.5	1707	9	PCT-US02-31357-39
18	18.8	85.5	1707	9	US-10-262-445-39
19	18.8	85.5	3237	11	US-10-170-235-5179
20	18.8	85.5	3237	11	US-60-452-680-612
21	18.8	85.5	3237	11	US-60-453-135-482
22	18.8	85.5	3445	11	US-60-453-050-482
					Sequence 3237, App

ALIGNMENTS

RESULT 1
US-09-270-437D-14
; Sequence 14, Application US/09270437D
; GENERAL INFORMATION:
; APPLICANT: Chen, Yao-Tseng
; APPLICANT: Gure, Ali
; APPLICANT: Tsang, Solam
; APPLICANT: Stockert, Elisabeth
; APPLICANT: Jager, Elke
; APPLICANT: Knuth, Alexander
; APPLICANT: Old, Lloyd J.
; TITLE OF INVENTION: Isolated Nucleic Acid Molecules Encoding Cancer Associated Antigen
; FILE REFERENCE: LUD 5538.1
; CURRENT APPLICATION NUMBER: US/09/270.437D
; CURRENT FILING DATE: 1999-03-16
; PRIOR APPLICATION NUMBER: 09/061,709
; PRIOR FILING DATE: 1998-04-17
; NUMBER OF SEQ ID NOS: 23
; SEQ ID NO 14
; LENGTH: 22
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
US-09-270-437D-14

Query Match 100.0%; Score 22; DB 5; Length 22;
Best Local Similarity 100.0%; Pred. No. 0.32; Indels 0; Gaps 0;
Matches 22; Conservative 0; Mismatches 0;
Qy 1 GACGTTGACACGGCGGTTTCT 22
Db 1 GACGTTGACACGGCGGTTTCT 22

RESULT 2
US-09-270-437D-8/c
; Sequence 8, Application US/09270437D
; GENERAL INFORMATION:
; APPLICANT: Chen, Yao-Tseng
; APPLICANT: Gure, Ali
; APPLICANT: Tsang, Solam
; APPLICANT: Stockert, Elisabeth
; APPLICANT: Jager, Elke
; APPLICANT: Knuth, Alexander

Sequence 1, Appli
Sequence 1, Appli
Sequence 31514, A
Sequence 43, Appli
Sequence 26, Appli
Sequence 31037, A
Sequence 34698, A
Sequence 997, App
Sequence 519026,
Sequence 35540, A
Sequence 38972, A
Sequence 38953, A
Sequence 34808, A
Sequence 44761, A
Sequence 6477, Ap
Sequence 7899, Ap
Sequence 63821, A
Sequence 45907, A
Sequence 14654, A
Sequence 46168, A
Sequence 8986, Ap
Sequence 44760, A
Sequence 46167, A

```

; APPLICANT: Old, Lloyd J.
; TITLE OF INVENTION: Isolated Nucleic Acid Molecules Encoding Cancer Associated Antigen
; FILE REFERENCE: LUD 5538.1
; CURRENT APPLICATION NUMBER: US/09/270,437D
; CURRENT FILING DATE: 1999-03-16
; PRIOR APPLICATION NUMBER: 09/061,709
; PRIOR FILING DATE: 1998-04-17
; NUMBER OF SEQ ID NOS: 23
; SEQ ID NO 8
; LENGTH: 3283
; TYPE: DNA
; ORGANISM: Homo sapiens
; NAME/KEY: CDS
; LOCATION: 3243
; OTHER INFORMATION: unsure of nucleotide
US-09-270-437D-8

Query Match      100.0%; Score 22; DB 5; Length 3283;
Best Local Similarity 100.0%; Pred. No. 0.51; Indels 0; Gaps 0;
Matches 22; Conservative 0; Mismatches 0;

Qy 1 GACGTTGACACGCGGTTTCT 22
Db 444 GACGTTGACACGCGGTTTCT 423

RESULT 3
US-09-270-437D-6/c
; Sequence 6, Application US/09270437D
; GENERAL INFORMATION:
; APPLICANT: Chen, Yao-Tseng
; APPLICANT: Gure, Ali
; APPLICANT: Tsang, Solam
; APPLICANT: Stockert, Elisabeth
; APPLICANT: Jager, Elke
; APPLICANT: Knuth, Alexander
; APPLICANT: Old, Lloyd J.
; TITLE OF INVENTION: Isolated Nucleic Acid Molecules Encoding Cancer Associated Antigen
; FILE REFERENCE: LUD 5538.1
; CURRENT APPLICATION NUMBER: US/09/270,437D
; CURRENT FILING DATE: 1999-03-16
; PRIOR APPLICATION NUMBER: 09/061,709
; PRIOR FILING DATE: 1998-04-17
; NUMBER OF SEQ ID NOS: 23
; SEQ ID NO 6
; LENGTH: 3412
; TYPE: DNA
; ORGANISM: Homo sapiens
; NAME/KEY: CDS
; LOCATION: 3372
; OTHER INFORMATION: unsure of nucleotide
US-09-270-437D-6

Query Match      100.0%; Score 22; DB 5; Length 3412;
Best Local Similarity 100.0%; Pred. No. 0.51; Indels 0; Gaps 0;
Matches 22; Conservative 0; Mismatches 0;

Qy 1 GACGTTGACACGCGGTTTCT 22
Db 444 GACGTTGACACGCGGTTTCT 423

RESULT 4
US-10-170-235-21870/c
; Sequence 21870, Application US/10170235
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig
; TITLE OF INVENTION: KITS, SUCH AS NUCLEIC ACID ARRAYS, COMPRISING A MAJORITY OF HUMAN
; TITLE OF INVENTION: TRANSCRIPTS, FOR DETECTING EXPRESSION AND OTHER USES THEREOF

```

```

; FILE REFERENCE: CL001380
; CURRENT APPLICATION NUMBER: US/10/170,235
; CURRENT FILING DATE: 2003-03-17
; NUMBER OF SEQ ID NOS: 42514
; SEQ ID NO 21870
; LENGTH: 3427
; TYPE: DNA
; ORGANISM: HUMAN
US-10-170-235-21870

Query Match      100.0%; Score 22; DB 8; Length 3427;
Best Local Similarity 100.0%; Pred. No. 0.52; Indels 0; Gaps 0;
Matches 22; Conservative 0; Mismatches 0;

Qy 1 GACGTTGACACGCGGTTTCT 22
Db 442 GACGTTGACACGCGGTTTCT 421

RESULT 5
US-60-453-135-3834/c
; Sequence 3834, Application US/60453135
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; APPLICANT: IAKOUBOVA, Olga
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001456
; CURRENT APPLICATION NUMBER: US/60/453,135
; CURRENT FILING DATE: 2003-03-10
; NUMBER OF SEQ ID NOS: 82762
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 3834
; LENGTH: 3427
; TYPE: DNA
; ORGANISM: Homo sapiens
US-60-453-135-3834

Query Match      100.0%; Score 22; DB 11; Length 3427;
Best Local Similarity 100.0%; Pred. No. 0.52; Indels 0; Gaps 0;
Matches 22; Conservative 0; Mismatches 0;

Qy 1 GACGTTGACACGCGGTTTCT 22
Db 442 GACGTTGACACGCGGTTTCT 421

RESULT 6
US-60-453-050-3834/c
; Sequence 3834, Application US/60453050
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; APPLICANT: LUKE, May
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: STENOSIS, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001457
; CURRENT APPLICATION NUMBER: US/60/453,050
; CURRENT FILING DATE: 2003-03-10
; NUMBER OF SEQ ID NOS: 82762
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 3834
; LENGTH: 3427
; TYPE: DNA
; ORGANISM: Homo sapiens
US-60-453-050-3834

Query Match      100.0%; Score 22; DB 11; Length 3427;
Best Local Similarity 100.0%; Pred. No. 0.52; Indels 0; Gaps 0;
Matches 22; Conservative 0; Mismatches 0;

Qy 1 GACGTTGACACGCGGTTTCT 22
Db 442 GACGTTGACACGCGGTTTCT 421

```

```
RESULT 7
US-10-170-235-22086/c
; Sequence 22086, Application US/10170235
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig
; TITLE OF INVENTION: KITS, SUCH AS NUCLEIC ACID ARRAYS, COMPRISING A MAJORITY OF HUMAN
; TITLE OF INVENTION: TRANSCRIPTS, FOR DETECTING EXPRESSION AND OTHER USES THEREOF
; FILE REFERENCE: CL001380
; CURRENT APPLICATION NUMBER: US/10/170,235
; CURRENT FILING DATE: 2003-03-17
; NUMBER OF SEQ ID NOS: 42514
; SEQ ID NO 22086
; LENGTH: 3475
; TYPE: DNA
; ORGANISM: HUMAN
US-10-170-235-22086

Query Match      100.0%; Score 22; DB 8; Length 3475;
Best Local Similarity 100.0%; Pred. No. 0.52;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACGTTGACAAACGCGGTTTCT 22
   |||||
Db 439 GACGTTGACAAACGCGGTTTCT 418

RESULT 8
US-60-453-135-3835/c
; Sequence 3835, Application US/60453135
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; APPLICANT: IAKUBOVA, Olga
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001456
; CURRENT APPLICATION NUMBER: US/60/453,135
; CURRENT FILING DATE: 2003-03-10
; NUMBER OF SEQ ID NOS: 82762
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3835
; LENGTH: 3475
; TYPE: DNA
; ORGANISM: Homo sapiens
US-60-453-135-3835

Query Match      100.0%; Score 22; DB 11; Length 3475;
Best Local Similarity 100.0%; Pred. No. 0.52;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACGTTGACAAACGCGGTTTCT 22
   |||||
Db 439 GACGTTGACAAACGCGGTTTCT 418

RESULT 9
US-60-453-050-3835/c
; Sequence 3835, Application US/60453050
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; APPLICANT: LUKE, May
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: STENOSIS, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001457
; CURRENT APPLICATION NUMBER: US/60/453,050
; CURRENT FILING DATE: 2003-03-10
; NUMBER OF SEQ ID NOS: 82762
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3835
; LENGTH: 3475
; TYPE: DNA
; ORGANISM: Homo sapiens
US-60-453-050-3835/c

Query Match      100.0%; Score 22; DB 11; Length 3475;
Best Local Similarity 100.0%; Pred. No. 0.52;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACGTTGACAAACGCGGTTTCT 22
   |||||
Db 439 GACGTTGACAAACGCGGTTTCT 418

RESULT 10
US-60-452-680-30351/c
; Sequence 30351, Application US/60452680
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; APPLICANT: GRUPE, Andrew
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: ALZHEIMER'S DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001450
; CURRENT APPLICATION NUMBER: US/60/452,680
; CURRENT FILING DATE: 2003-03-07
; NUMBER OF SEQ ID NOS: 116213
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 30351
; LENGTH: 201
; TYPE: DNA
; ORGANISM: Homo sapiens
US-60-452-680-30351

Query Match      85.5%; Score 18.8; DB 11; Length 201;
Best Local Similarity 90.9%; Pred. No. 15;
Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GACGTTGACAAACGCGGTTTCT 22
   |||||
Db 135 GACGTTGACAAACGCGGTTTCT 114

RESULT 11
US-60-452-680-30361/c
; Sequence 30361, Application US/60452680
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; APPLICANT: GRUPE, Andrew
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: ALZHEIMER'S DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001450
; CURRENT APPLICATION NUMBER: US/60/452,680
; CURRENT FILING DATE: 2003-03-07
; NUMBER OF SEQ ID NOS: 116213
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 30361
; LENGTH: 201
; TYPE: DNA
; ORGANISM: Homo sapiens
US-60-452-680-30361

Query Match      85.5%; Score 18.8; DB 11; Length 201;
Best Local Similarity 90.9%; Pred. No. 15;
Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GACGTTGACAAACGCGGTTTCT 22
   |||||
Db 171 GACGTTGACAAACGCGGTTTCT 150

RESULT 12
US-60-453-135-20375/c
; Sequence 20375, Application US/60453135
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; APPLICANT: IAKUBOVA, Olga
```

```
US-60-453-050-3835

Query Match      100.0%; Score 22; DB 11; Length 3475;
Best Local Similarity 100.0%; Pred. No. 0.52;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACGTTGACAAACGCGGTTTCT 22
   |||||
Db 439 GACGTTGACAAACGCGGTTTCT 418

RESULT 10
US-60-452-680-30351/c
; Sequence 30351, Application US/60452680
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; APPLICANT: GRUPE, Andrew
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: ALZHEIMER'S DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001450
; CURRENT APPLICATION NUMBER: US/60/452,680
; CURRENT FILING DATE: 2003-03-07
; NUMBER OF SEQ ID NOS: 116213
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 30351
; LENGTH: 201
; TYPE: DNA
; ORGANISM: Homo sapiens
US-60-452-680-30351

Query Match      85.5%; Score 18.8; DB 11; Length 201;
Best Local Similarity 90.9%; Pred. No. 15;
Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GACGTTGACAAACGCGGTTTCT 22
   |||||
Db 135 GACGTTGACAAACGCGGTTTCT 114

RESULT 11
US-60-452-680-30361/c
; Sequence 30361, Application US/60452680
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; APPLICANT: GRUPE, Andrew
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: ALZHEIMER'S DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001450
; CURRENT APPLICATION NUMBER: US/60/452,680
; CURRENT FILING DATE: 2003-03-07
; NUMBER OF SEQ ID NOS: 116213
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 30361
; LENGTH: 201
; TYPE: DNA
; ORGANISM: Homo sapiens
US-60-452-680-30361

Query Match      85.5%; Score 18.8; DB 11; Length 201;
Best Local Similarity 90.9%; Pred. No. 15;
Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GACGTTGACAAACGCGGTTTCT 22
   |||||
Db 171 GACGTTGACAAACGCGGTTTCT 150

RESULT 12
US-60-453-135-20375/c
; Sequence 20375, Application US/60453135
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; APPLICANT: IAKUBOVA, Olga
```

us-09-270-437d-14.rnpn

Mon Apr 21 10:30:44 2003

; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
 ; TITLE OF INVENTION: MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF
 ; FILE REFERENCE: CL001456
 ; CURRENT APPLICATION NUMBER: US/60/453,135
 ; CURRENT FILING DATE: 2003-03-10
 ; NUMBER OF SEQ ID NOS: 82762
 ; SOFTWARE: FastSEQ for Windows Version 4.0
 ; SEQ ID NO 20375
 ; LENGTH: 201
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; US-60-453-135-20375

Query Match 85.5%; Score 18.8; DB 11; Length 201;
 Best Local Similarity 90.9%; Pred. No. 15;
 Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GACGTTGACAAACGCGGTTTCT 22
 |||||
 Db 135 GACGTTGACAAACAGTGTTTCT 114

RESULT 13
 US-60-453-135-20385/c
 ; Sequence 20385, Application US/60453135
 ; GENERAL INFORMATION:
 ; APPLICANT: CARGILL, Michele
 ; APPLICANT: IAKOUBOVA, Olga
 ; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
 ; TITLE OF INVENTION: MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF
 ; FILE REFERENCE: CL001456
 ; CURRENT APPLICATION NUMBER: US/60/453,135
 ; CURRENT FILING DATE: 2003-03-10
 ; NUMBER OF SEQ ID NOS: 82762
 ; SOFTWARE: FastSEQ for Windows Version 4.0
 ; SEQ ID NO 20385
 ; LENGTH: 201
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; US-60-453-135-20385

Query Match 85.5%; Score 18.8; DB 11; Length 201;
 Best Local Similarity 90.9%; Pred. No. 15;
 Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GACGTTGACAAACGCGGTTTCT 22
 |||||
 Db 171 GACGTTGACAAACAGTGTTTCT 150

RESULT 14
 US-60-453-050-20375/c
 ; Sequence 20375, Application US/60453050
 ; GENERAL INFORMATION:
 ; APPLICANT: CARGILL, Michele
 ; APPLICANT: LUKE, May
 ; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
 ; TITLE OF INVENTION: STENOSIS, METHODS OF DETECTION AND USES THEREOF
 ; FILE REFERENCE: CL001457
 ; CURRENT APPLICATION NUMBER: US/60/453,050
 ; CURRENT FILING DATE: 2003-03-10
 ; NUMBER OF SEQ ID NOS: 82762
 ; SOFTWARE: FastSEQ for Windows Version 4.0
 ; SEQ ID NO 20375
 ; LENGTH: 201
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; US-60-453-050-20375

Query Match 85.5%; Score 18.8; DB 11; Length 201;
 Best Local Similarity 90.9%; Pred. No. 15;
 Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GACGTTGACAAACGCGGTTTCT 22
 |||||
 Db 135 GACGTTGACAAACAGTGTTTCT 114

RESULT 15
 US-60-453-050-20385/c
 ; Sequence 20385, Application US/60453050
 ; GENERAL INFORMATION:
 ; APPLICANT: CARGILL, Michele
 ; APPLICANT: LUKE, May
 ; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
 ; TITLE OF INVENTION: STENOSIS, METHODS OF DETECTION AND USES THEREOF
 ; FILE REFERENCE: CL001457
 ; CURRENT APPLICATION NUMBER: US/60/453,050
 ; CURRENT FILING DATE: 2003-03-10
 ; NUMBER OF SEQ ID NOS: 82762
 ; SOFTWARE: FastSEQ for Windows Version 4.0
 ; SEQ ID NO 20385
 ; LENGTH: 201
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; US-60-453-050-20385

Query Match 85.5%; Score 18.8; DB 11; Length 201;
 Best Local Similarity 90.9%; Pred. No. 15;
 Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GACGTTGACAAACGCGGTTTCT 22
 |||||
 Db 171 GACGTTGACAAACAGTGTTTCT 150

Search completed: April 18, 2003, 10:50:34
 Job time : 248.167 secs

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 18, 2003, 04:47:40 ; Search time 766 Seconds
(without alignments)
465.145 Million cell updates/sec

Title: US-09-270-437D-14

Perfect score: 22

Sequence: 1 gacgttgacacgcgggtttct 22

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST.*

1: em_estba.*

2: em_esthum.*

3: em_estin.*

4: em_estmu.*

5: em_estov.*

6: em_estpl.*

7: em_estro.*

8: em_htc.*

9: gb_est1.*

10: gb_est2.*

11: gb_htc.*

12: gb_est3.*

13: gb_est4.*

14: gb_est5.*

15: em_estfun.*

16: em_estom.*

17: gb_gss.*

18: em_gss_hum.*

19: em_gss_inv.*

20: em_gss_pln.*

21: em_gss_vrt.*

22: em_gss_fun.*

23: em_gss_mam.*

24: em_gss_mus.*

25: em_gss_other.*

26: em_gss_pro.*

27: em_gss_rod.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	22	100.0	402	12	BG003431
C 2	22	100.0	453	9	AL121466
C 3	22	100.0	531	12	BG009869
C 4	22	100.0	594	13	BI858380
C 5	22	100.0	620	12	BF914566
C 6	22	100.0	642	17	AQ479871

7	22	100.0	688	17	AG087457
8	22	100.0	724	14	BM977219
9	22	100.0	759	10	BE622021
10	22	100.0	851	14	BQ221568
11	22	100.0	856	12	BG480515
12	22	100.0	889	12	BG748346
13	22	100.0	1089	14	BM806132
14	21	95.5	682	17	AG048517
15	20.4	92.7	448	10	BB852457
16	20.4	92.7	458	10	BB852318
17	20.4	92.7	462	10	BB860142
18	20.4	92.7	469	10	BB853640
19	20.4	92.7	541	14	BQ770242
20	20.4	92.7	629	13	BI655144
21	20.4	92.7	663	10	BB643209
22	20.4	92.7	668	10	BB618315
23	20.4	92.7	815	13	BI688088
24	20.4	92.7	905	13	BI654462
25	19.4	88.2	464	10	BB862701
26	18.8	85.5	489	10	BB855310
27	17.8	80.9	522	12	BG279958
28	17.8	80.9	790	12	BF140399
29	17.8	80.9	1120	17	CNS06502
30	17.2	78.2	329	17	AQ091026
31	17.2	78.2	479	14	BU028477
32	17.2	78.2	511	9	AI960673
33	17.2	78.2	517	12	BG543604
34	17.2	78.2	522	17	AQ467039
35	17.2	78.2	543	14	BU026166
36	17.2	78.2	548	17	AZ178687
37	17.2	78.2	555	14	BQ488142
38	17.2	78.2	562	13	EM186239
39	17.2	78.2	576	14	BU024011
40	17.2	78.2	581	14	BU025444
41	17.2	78.2	581	14	BU029203
42	17.2	78.2	616	17	AZ792519
43	17.2	78.2	625	14	BU025074
44	17.2	78.2	637	10	BE263528
45	17.2	78.2	640	14	BU027309

ALIGNMENTS

RESULT 1
BG003431/c
LOCUS RCI-GN0198-171100-021-f07 GN0198 Homo sapiens cDNA, mRNA sequence.
DEFINITION
ACCESSION BG003431
VERSION
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 402)

402 bp mRNA linear EST 24-JAN-2001
RCI-GN0198-171100-021-f07 GN0198 Homo sapiens cDNA, mRNA sequence.

BG003431.1 GI:12443763

DIAS Neto, E., Garcia Correa, R., Verjovskii-Almeida, S., Briones, M.R.,
Negai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F.,
Goldman, G.H., Carvalho, A.F., Matsukuma, A., Baia, G.S., Simpson, D.H.,
Brunstein, A., deOliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare,
M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and
Simpson, A.J.
Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
20202663
Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001

Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?t1=RC1&t2=RC1-GN0198-171100-021-f07&t3=2000-11-17&t4=1)
Seq primer: puc 18 forward
High quality sequence start: 15
High quality sequence stop: 400.

FEATURES
source
1..402
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="GN0198"
/dev_stage="Adult"
/note="Organ: placenta normal; Vector: puc18; Site_1: SmaI
; Site_2: SmaI; A mini-library was made by cloning
products derived from ORESTES PCR (U.S. Letters Patent
application No. 196,716 - Ludwig Institute for Cancer
Research) profiles into the pUC 18 vector. Reverse
transcription of tissue mRNA and cDNA amplification were
performed under low stringency conditions."

BASE COUNT 117 a 98 c 109 g 78 t

ORIGIN
Query Match 100.0%; Score 22; DB 12; Length 402;
Best Local Similarity 100.0%; Pred. No. 4.2; Indels 0; Gaps 0;
Matches 22; Conservative 0; Mismatches 0

QY 1 GACGTTGACACGGCGGTTTCT 22
|||||
Db 195 GACGTTGACACGGCGGTTTCT 174

RESULT 2
AL121466/c 453 bp mRNA linear EST 25-FEB-2000
LOCUS
DEFINITION DKFZp762L097_r1.762 (synonym: hmel2) Homo sapiens cDNA clone
AL121466
VERSION DKFZp762L097.5', mRNA sequence.
AL121466.1 GI:5927467
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 453)
AUTHORS Ottenwaelder,B., Obermaier,B., Mewes,H.W., Gassenhuber,J. and
Wiemann,S.
TITLE EST (Ottenwaelder, et al.)
JOURNAL Unpublished (1999)
COMMENT Contact: Ottenwaelder B
MIPS
Am Klopferspitz 18a D-82152 Martinsried, Germany
This is the 5' sequence of the clone insert
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
sequenced by Medigenomix (Martinsried/Germany) within the cDNA
sequencing consortium of the German Genome Project. No s1 sequence
available.
This clone (DKFZp762L097) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.

FEATURES
source
1..453
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="DKFZp762L097"
/clone_lib="762 (synonym: hmel2)"
/tissue_type="melanoma (Mewo cell line)"
/dev_stage="adult"
/lab_host="PH108"
/note="Vector: pSPort1; Site_1: NotI; Site_2: SalI"

BASE COUNT 117 a 121 c 137 g 78 t

ORIGIN

Query Match 100.0%; Score 22; DB 9; Length 453;
Best Local Similarity 100.0%; Pred. No. 4.5; Indels 0; Gaps 0;
Matches 22; Conservative 0; Mismatches 0

QY 1 GACGTTGACACGGCGGTTTCT 22
|||||
Db 430 GACGTTGACACGGCGGTTTCT 409

RESULT 3

BG009869 531 bp mRNA linear EST 24-JAN-2001
LOCUS
DEFINITION QV1-GN0320-041200-532-c06 GN0320 Homo sapiens cDNA, mRNA sequence.
ACCESSION BG009869
VERSION BG009869.1 GI:12456498
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 531)
AUTHORS Dias Neto,E., Garcia Correa,R., Vertovski-Almeida,S., Briones,M.R.,
Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
Brunstein,A., deOliveira,P., Bucher,P., Jongeneel,C.V., O'Hare
,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
Simpson,A.J.
TITLE Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE 20202663
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?t1=QV1&t2=QV1-GN0320-
041200-532-c06&t3=2000-12-04&t4=1)
Seq primer: puc 18 forward
High quality sequence stop: 531.
Location/Qualifiers
1..531
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="GN0320"
/dev_stage="Adult"

FEATURES

source
1..531
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="GN0320"
/dev_stage="Adult"
/note="Organ: placenta normal; Vector: puc18; Site_1: SmaI
; Site_2: SmaI; A mini-library was made by cloning
products derived from ORESTES PCR (U.S. Letters Patent
application No. 196,716 - Ludwig Institute for Cancer
Research) profiles into the pUC 18 vector. Reverse
transcription of tissue mRNA and cDNA amplification were
performed under low stringency conditions."

BASE COUNT 104 a 138 c 135 g 154 t

ORIGIN
Query Match 100.0%; Score 22; DB 12; Length 531;
Best Local Similarity 100.0%; Pred. No. 4.9; Indels 0; Gaps 0;
Matches 22; Conservative 0; Mismatches 0

QY 1 GACGTTGACACGGCGGTTTCT 22
|||||
Db 339 GACGTTGACACGGCGGTTTCT 360

RESULT 4

B1858380/c
LOCUS B1858380 594 bp mRNA linear EST 10-OCT-2001
DEFINITION 603383965F1 NIH_MGC_87 Homo sapiens cDNA clone IMAGE:5392792 5', mRNA sequence.
ACCESSION B1858380
VERSION B1858380.1 GI:15999127
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC) Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-romail.nih.gov
 Tissue procurement: DCTD/DTF
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Plate: LLM12001 row: a column: 17
 High quality sequence stop: 594.
 Location/Qualifiers
 1..594
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:5392792"
 /tissue_type="mammary adenocarcinoma, cell line"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: breast; Vector: pCMV-SPORT6; Site 1: NotI; Site 2: SalI; Cloned unidirectionally; oligo-dr primed. Average insert size 1.383 Kb. Library enriched for full-length clones and constructed by Life Technologies. Note: this is a NIH MGC Library."
 167 a 157 c 150 g 120 t
 BASE COUNT 167 a 157 c 150 g 120 t
 ORIGIN
 Query Match 100.0%; Score 22; DB 13; Length 594;
 Best Local Similarity 100.0%; Pred. No. 5.2;
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 GACGTTGACACGCGGGTTTCT 22
 |||||
 Db 197 GACGTTGACACGCGGGTTTCT 176
 |||||
 RESULT 5
 BF914566/c
LOCUS BF914566 620 bp mRNA linear EST 18-JAN-2001
DEFINITION IL3-UT0114-011200-362-E06 UT0114 Homo sapiens cDNA, mRNA sequence.
ACCESSION BF914566
VERSION BF914566.1 GI:12306024
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS 1 (bases 1 to 620)
 Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R., Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F., Goldman, G.H., Carvalho, A.F., Matsukuma, A., Baia, G.S., Simpson, D.H., Brunstein, A., deOliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and Simpson, A.J.J.
TITLE Shotgun sequencing of the human transcriptome with ORF expressed sequence tags
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE 20202663

COMMENT Contact: Simpson A.J.J.
 Laboratory of Cancer Genetics
 Ludwig Institute for Cancer Research
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil
 Tel: +55-11-2704922
 Fax: +55-11-2707001
 Email: asimpson@ludwig.org.br
 This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL
 (<http://www.ludwig.org.br/scripts/gethtml2.pl?tl=IL3&2=IL3-UT0114-011200-362-E06&t3=2000-12-01&t4=1>)
 Seq primer: puc 18 forward
 High quality sequence stop: 619.
FEATURES Location/Qualifiers
 1..620
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_lib="UT0114"
 /dev_stage="Adult"
 /note="Organ: uterus tumor; Vector: puc18; Site 1: SmaI; Site 2: SmaI; A mini-library was made by cloning products derived from OREGRES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."
 181 a 157 c 150 g 132 t
 BASE COUNT 181 a 157 c 150 g 132 t
 ORIGIN
 Query Match 100.0%; Score 22; DB 12; Length 620;
 Best Local Similarity 100.0%; Pred. No. 5.3;
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 GACGTTGACACGCGGGTTTCT 22
 |||||
 Db 385 GACGTTGACACGCGGGTTTCT 364
 |||||
 RESULT 6
 AQ479871
LOCUS AQ479871 642 bp DNA linear GSS 23-APR-1999
DEFINITION RPCI-11-274C8.TV RPCI-11 Homo sapiens genomic clone RPCI-11-274C8, DNA sequence.
ACCESSION AQ479871
VERSION AQ479871.1 GI:4661990
KEYWORDS GSS.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS 1 (bases 1 to 642)
 Zhao, S., Adams, M.D., Nierman, W., Malek, J., de Jong, P. and Venter, J.C.
TITLE Use of BAC End Sequences from Library RPCI-11 for Sequence-Ready Map Building
JOURNAL Unpublished (1997)
COMMENT Contact: Shaying Zhao, William Nierman, Mark Adams
 Department of Eukaryotic Genomics
 The Institute for Genomic Research
 9712 Medical Center Dr., Rockville, MD 20850
 Tel: 301 838 0200
 Fax: 301 838 0208
 Email: hbe@igr.org
 Clones are derived from the human BAC library RPCI-11. For BAC library availability, please contact Pieter de Jong (pieter@dejong.med.buffalo.edu). Clones may be purchased from BACPAC Resources (<http://bacpac.med.buffalo.edu/ordering>) or from Research Genet ca (info@resgen.com). BAC end search page: http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html.
 Seq primer: T7
 Class: BAC ends.
FEATURES Location/Qualifiers

```

source
1. .642
/organism="Homo sapiens"
/db_xref="GDB:7604887"
/db_xref="taxon:9606"
/clone="RPCI-11-274C8"
/clone_lib="RPCI-11"
/sex="Male"
/cell_type="Lymphocytes"
/notes="Vector: pBACe3.6; Site 1: EcoRI; Site 2: EcoRI;
RPC11 Human Male BAC Library"
BASE COUNT 207 a 117 c 95 g 223 t
ORIGIN
Query Match 100.0%; Score 22; DB 17; Length 642;
Best Local Similarity 100.0%; Pred. No. 5.4;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GACGTTGACACGCGGTTTCT 22
|||||
Db 76 GACGTTGACACGCGGTTTCT 97

RESULT 7
AG087457 688 bp DNA linear GSS 03-NOV-2001
LOCUS Pan troglodytes DNA, clone: PTB-086B21.F, genomic survey sequence.
ACCESSION AG087457
VERSION AG087457.1 GI:16639259
KEYWORDS GSS.
SOURCE Pan troglodytes male lymphoblast DNA, clone_lib:PTB Chimpanzee Male
BAC Library clone:PTB-086B21.F.
ORGANISM
Pan troglodytes
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
REFERENCE
1 Fujiyama,A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T.,
Totoki,Y., Watanabe,H. and Sakaki,Y.
BAC end sequences of Library PTB
Unpublished
2 (bases 1 to 688)
Fujiyama,A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T.,
Totoki,Y., Watanabe,H. and Sakaki,Y.
Direct Submission
Submitted (02-AUG-2001) Asao Fujiyama, The Institute of Physical
and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
1-7-22 Suehiro-chou,Tsukumi-ku, Yokohama, Kanagawa 230-0045, Japan
(E-mail:chimbeg@gs.riken.go.jp, URL:http://hgsc.riken.go.jp/,
Tel:81-45-503-9111, Fax:81-45-503-9170)
Clones are derived from the chimpanzee BAC library PTB This BAC end
was generated during the Rad process and may have higher chance of
clone tracking errors.
PRIMERS
Sequencing: -21M13
LIBRARY
Vector : pKS145
R.Site 1 : SacI
R.Site 2 : SacI
Location/Qualifiers
1. .688
/organism="Pan troglodytes"
/db_xref="taxon:9598"
/clone="PTB-086B21.F"
/sex="male"
/cell_type="lymphoblast"
/clone_lib="PTB Chimpanzee Male BAC Library"
BASE COUNT 217 a 153 c 120 g 198 t
ORIGIN
Query Match 100.0%; Score 22; DB 17; Length 688;
Best Local Similarity 100.0%; Pred. No. 5.6;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GACGTTGACACGCGGTTTCT 22
|||||
Db 76 GACGTTGACACGCGGTTTCT 97

source
1. .642
/organism="Homo sapiens"
/db_xref="GDB:7604887"
/db_xref="taxon:9606"
/clone="RPCI-11-274C8"
/clone_lib="RPCI-11"
/sex="Male"
/cell_type="Lymphocytes"
/notes="Vector: pBACe3.6; Site 1: EcoRI; Site 2: EcoRI;
RPC11 Human Male BAC Library"
BASE COUNT 207 a 117 c 95 g 223 t
ORIGIN
Query Match 100.0%; Score 22; DB 17; Length 642;
Best Local Similarity 100.0%; Pred. No. 5.4;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GACGTTGACACGCGGTTTCT 22
|||||
Db 76 GACGTTGACACGCGGTTTCT 97

RESULT 8
BM977219 724 bp mRNA linear EST 21-MAR-2002
LOCUS UI-CF-DUI-ads-n-12-0-UI.s1 UI-CF-DUI Homo sapiens cDNA clone
DEFINITION UI-CF-DUI-ads-n-12-0-UI 3', mRNA sequence.
ACCESSION BM977219
VERSION BM977219.1 GI:19595413
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 724)
Bonaldo,M.F., Lennon,G. and Soares,M.B.
Normalization and subtraction: two approaches to facilitate gene
discovery
Genome Res. 6 (9), 791-806 (1996)
JOURNAL
MEDLINE
COMMENT
Contact: McCray, PB
9704447
McCray Lab
University of Iowa
2024 University of Iowa Med Labs, Iowa City, IA 52242, USA
Tel: 319 356 4866
Fax: 319 356 7171
Email: paul-mccray@uiowa.edu
Tissue Procurement: Dr. M. J. Welsh, University of Iowa
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Researchers may obtain clones from Research
Genetics (www.resgen.com).
Seq primer: M13 FORWARD
POLYA=Yes.
Location/Qualifiers
1. .724
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="UI-CF-DUI-ads-n-12-0-UI"
/clone_lib="UI-CF-DUI"
/tissue_type="Primary Lung Epithelial Cells"
/dev_stage="Adult"
/lab_host="DH10B (Life Technologies) (T1 phage resistant)"
/notes="Organ: Lung; Vector: pVT3-Pac (Pharmacia) with a
modified polylinker; Site 1: EcoR I; Site 2: Not I;
UI-CF-DUI is a normalized cDNA library containing the
following tissue(s): Primary Lung Epithelial Cells The
library was constructed according to Bonaldo, Lennon and
Soares, Genome Research, 6:791-806, 1996. First strand
cDNA synthesis was primed with an oligo-dT primer
containing a Not I site. Double stranded cDNA was ligated
to an EcoR I adaptor, digested with Not I, and cloned
directionally into pVT3-Pac vector. The oligonucleotide
used to prime the synthesis of first-strand cDNA contains
a library tag sequence that is located between the Not I
site and the (dT)18 tail. The sequence tag for this
library is GCGTGTAGGC.
TAG LIB=UI-CF-DUI
TAG TISSUE=Lung Epithelial Cells Tissue nos 359-368
TAG_SEQ=GCGTGTAGGC"
BASE COUNT 219 a 174 c 182 g 149 t
ORIGIN
Query Match 100.0%; Score 22; DB 14; Length 724;
Best Local Similarity 100.0%; Pred. No. 5.8;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GACGTTGACACGCGGTTTCT 22
|||||
Db 87 GACGTTGACACGCGGTTTCT 66

```

```

RESULT 9
BB622021/c
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
BB622021
601440873F1 NIH_MGC_72 Homo sapiens cDNA clone IMAGE:3915546 5',
mRNA sequence.
BB622021
BB622021.1 GI:9892961
human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NIH-MGC http://mgs.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC/DCTD/DTP
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
CDNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM9739 row: a column: 19
High quality sequence stop: 622.
Location/Qualifiers
1..759
/organism="Homo sapiens"
/db xref="taxon:9606"
/clone="IMAGE:3915546"
/clone_lib="NIH_MGC_72"
/tissue_type="melanotic melanoma"
/lab_host="DH10B (phage-resistant)"
/notes="Organ: skin; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 2 kb. Library constructed by Life
Technologies."
BASE COUNT 228 a 209 c 188 g 134 t
ORIGIN
1..759
GACGTTGACAAACGGCGTTTCT 22
|||||
GACGTTGACAAACGGCGTTTCT 143

Query Match 100.0%; Score 22; DB 10; Length 759;
Best Local Similarity 100.0%; Pred. No. 5.9;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACGTTGACAAACGGCGTTTCT 22
|||||
GACGTTGACAAACGGCGTTTCT 143
DB 164 GACGTTGACAAACGGCGTTTCT 143

RESULT 10
BQ221568/c
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
BQ221568
AGENCOURT 7559207 NIH_MGC_72 Homo sapiens cDNA clone IMAGE:6046477
5', mRNA sequence.
BQ221568
BQ221568.1 GI:20402968
human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NIH-MGC http://mgs.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC/DCTD/DTP
CDNA Library Preparation: Life Technologies, Inc.

```

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM13292 row: f column: 14
High quality sequence stop: 637.
Location/Qualifiers
1..851
/organism="Homo sapiens"
/db xref="taxon:9606"
/clone="IMAGE:6046477"
/clone_lib="NIH_MGC_72"
/tissue_type="melanotic melanoma"
/lab_host="DH10B (phage-resistant)"
/notes="Organ: skin; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 2 kb. Library constructed by Life
Technologies."

BASE COUNT 229 a 234 c 231 g 157 t
ORIGIN

Query Match 100.0%; Score 22; DB 14; Length 851;
Best Local Similarity 100.0%; Pred. No. 6.3;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACGTTGACAAACGGCGTTTCT 22
|||||
DB 398 GACGTTGACAAACGGCGTTTCT 377

RESULT 11
BG480515/c
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

BG480515
602529520F1 NIH_MGC_21 Homo sapiens cDNA clone IMAGE:4652981 5',
mRNA sequence.
BG480515
BG480515.1 GI:13412705
EST.
human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 856)
NIH-MGC http://mgs.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCM1438 row: h column: 06
High quality sequence stop: 661.
Location/Qualifiers
1..856
/organism="Homo sapiens"
/db xref="taxon:9606"
/clone="IMAGE:4652981"
/clone_lib="NIH_MGC_21"
/tissue_type="choriocarcinoma"
/lab_host="DH10B (phage-resistant)"
/notes="Organ: placenta; Vector: pOTB7; Site 1: XhoI;
Site 2: EcoRI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCACGAG(G). Size-selected >500bp
for average insert size 1.8kb. Library constructed by
Ling Hong in the laboratory of Gerald M. Rubin (University
of California, Berkeley) using ZAP-cDNA synthesis kit

FEATURES
source

<p>BM806132 EST. GI:19122955</p> <p>human.</p> <p>Homo sapiens</p> <p>Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.</p> <p>1 (bases 1 to 1089)</p> <p>NIH-MGC http://mgc.nci.nih.gov/</p> <p>National Institutes of Health, Mammalian Gene Collection (MGC)</p> <p>Unpublished (1999)</p> <p>Contact: Robert Strausberg, Ph.D.</p> <p>Email: cgabbs@mail.nih.gov</p> <p>Tissue Procurement: ATCC</p> <p>CDNA Library Preparation: Life Technologies, Inc.</p> <p>CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)</p> <p>DNA Sequencing by: Agencourt Bioscience Corporation</p> <p>Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov</p> <p>Plate: LLAM12276 row: c column: 13</p> <p>High quality sequence stop: 688.</p>	<p>BM806132 EST. GI:19122955</p> <p>human.</p> <p>Homo sapiens</p> <p>Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.</p> <p>1 (bases 1 to 1089)</p> <p>NIH-MGC http://mgc.nci.nih.gov/</p> <p>National Institutes of Health, Mammalian Gene Collection (MGC)</p> <p>Unpublished (1999)</p> <p>Contact: Robert Strausberg, Ph.D.</p> <p>Email: cgabbs@mail.nih.gov</p> <p>Tissue Procurement: ATCC</p> <p>CDNA Library Preparation: Life Technologies, Inc.</p> <p>CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)</p> <p>DNA Sequencing by: Agencourt Bioscience Corporation</p> <p>Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov</p> <p>Plate: LLAM12276 row: c column: 13</p> <p>High quality sequence stop: 688.</p>
<p>Accession Version Keywords Source</p> <p>ORGANISM</p> <p>REFERENCE</p> <p>AUTHORS</p> <p>TITLE</p> <p>JOURNAL</p> <p>COMMENT</p>	<p>Accession Version Keywords Source</p> <p>ORGANISM</p> <p>REFERENCE</p> <p>AUTHORS</p> <p>TITLE</p> <p>JOURNAL</p> <p>COMMENT</p>
<p>Query Match 100.0%; Score 22; DB 12; Length 856; Best Local Similarity 100.0%; Pred. No. 6.3; Indels 0; Gaps 0; Matches 22; Conservative 0; Mismatches 0</p> <p>QY 1 GACGTTGACACGGCGGTTTCT 22 Db 452 GACGTTGACACGGCGGTTTCT 431 </p>	<p>Query Match 100.0%; Score 22; DB 12; Length 856; Best Local Similarity 100.0%; Pred. No. 6.3; Indels 0; Gaps 0; Matches 22; Conservative 0; Mismatches 0</p> <p>QY 1 GACGTTGACACGGCGGTTTCT 22 Db 452 GACGTTGACACGGCGGTTTCT 431 </p>
<p>RESULT 12 BG748346/c LOCUS DEFINITION ACCESSION VERSION KEYWORDS SOURCE</p> <p>ORGANISM</p> <p>REFERENCE</p> <p>AUTHORS</p> <p>TITLE</p> <p>JOURNAL</p> <p>COMMENT</p>	<p>RESULT 12 BG748346/c LOCUS DEFINITION ACCESSION VERSION KEYWORDS SOURCE</p> <p>ORGANISM</p> <p>REFERENCE</p> <p>AUTHORS</p> <p>TITLE</p> <p>JOURNAL</p> <p>COMMENT</p>
<p>Human.</p> <p>Homo sapiens</p> <p>Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.</p> <p>1 (bases 1 to 889)</p> <p>NIH-MGC http://mgc.nci.nih.gov/</p> <p>National Institutes of Health, Mammalian Gene Collection (MGC)</p> <p>Unpublished (1999)</p> <p>Contact: Robert Strausberg, Ph.D.</p> <p>Email: cgabbs@mail.nih.gov</p> <p>Tissue Procurement: ATCC</p> <p>CDNA Library Preparation: Ling Hong/Rubin Laboratory</p> <p>CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)</p> <p>DNA Sequencing by: Incyte Genomics, Inc.</p> <p>Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov</p> <p>Plate: LLCM1677 row: c column: 02</p> <p>High quality sequence stop: 773.</p>	<p>Human.</p> <p>Homo sapiens</p> <p>Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.</p> <p>1 (bases 1 to 889)</p> <p>NIH-MGC http://mgc.nci.nih.gov/</p> <p>National Institutes of Health, Mammalian Gene Collection (MGC)</p> <p>Unpublished (1999)</p> <p>Contact: Robert Strausberg, Ph.D.</p> <p>Email: cgabbs@mail.nih.gov</p> <p>Tissue Procurement: ATCC</p> <p>CDNA Library Preparation: Ling Hong/Rubin Laboratory</p> <p>CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)</p> <p>DNA Sequencing by: Incyte Genomics, Inc.</p> <p>Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov</p> <p>Plate: LLCM1677 row: c column: 02</p> <p>High quality sequence stop: 773.</p>
<p>Features</p> <p>source</p> <p>1..889</p> <p>/organism="Homo sapiens"</p> <p>/db xref="taxon:9606"</p> <p>/clone lib="NIH MGC 43"</p> <p>/tissue type="normal pigmented retinal epithelium"</p> <p>/lab host="DH10B (phage-resistant)"</p> <p>/note="Organ: eye; Vector: pOTB7; Site 1: XhoI; Site 2: EcoRI; CDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGACGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-CDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH MGC library. "</p>	<p>Features</p> <p>source</p> <p>1..889</p> <p>/organism="Homo sapiens"</p> <p>/db xref="taxon:9606"</p> <p>/clone lib="NIH MGC 43"</p> <p>/tissue type="normal pigmented retinal epithelium"</p> <p>/lab host="DH10B (phage-resistant)"</p> <p>/note="Organ: eye; Vector: pOTB7; Site 1: XhoI; Site 2: EcoRI; CDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGACGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-CDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH MGC library. "</p>
<p>BASE COUNT 242 a 241 c 235 g 171 t</p> <p>ORIGIN</p> <p>Query Match 100.0%; Score 22; DB 12; Length 889; Best Local Similarity 100.0%; Pred. No. 6.5; Indels 0; Gaps 0; Matches 22; Conservative 0; Mismatches 0</p> <p>QY 1 GACGTTGACACGGCGGTTTCT 22 Db 374 GACGTTGACACGGCGGTTTCT 353 </p>	<p>BASE COUNT 242 a 241 c 235 g 171 t</p> <p>ORIGIN</p> <p>Query Match 100.0%; Score 22; DB 12; Length 889; Best Local Similarity 100.0%; Pred. No. 6.5; Indels 0; Gaps 0; Matches 22; Conservative 0; Mismatches 0</p> <p>QY 1 GACGTTGACACGGCGGTTTCT 22 Db 374 GACGTTGACACGGCGGTTTCT 353 </p>
<p>RESULT 13 BM806132/c LOCUS DEFINITION</p> <p>AGENCOURT_6553922 NIH_MGC_71 Homo sapiens cDNA clone IMAGE:5555652</p>	<p>RESULT 13 BM806132/c LOCUS DEFINITION</p> <p>AGENCOURT_6553922 NIH_MGC_71 Homo sapiens cDNA clone IMAGE:5555652</p>

clone tracking errors.

PRIMERS

Sequencing: -21M13

LIBRARY

Vector : PKS145

R.Site 1 : SacI

R.Site 2 : SacI

Location/Qualifiers

1. .682

/organism="Pan troglodytes"

/db_xref="taxon:9598"

/clone="PTB-028F21.F"

/sex="male"

/cell_type="lymphoblast"

/clone_lib="PTB Chimpanzee Male BAC Library"

217 a 147 c 112 g 199 t 7 others

BASE COUNT

ORIGIN

Query Match 95.5%; Score 21; DB 17; Length 682;

Best Local Similarity 95.5%; Pred. No. 17;

Matches 21; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GAGCTTGACACGGCGGTTTCT 22

DB 255 GAGCTTGACACGGCGGTTTCT 276

RESULT 15

BB852457/c

LOCUS

DEFINITION

CDNA clone G370007124 5', mRNA sequence.

BB852457

BB852457.1

GI:17093911

EST.

house mouse.

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 448)

Akimura,T., Arakawa,T., Carninci,P., Furuno,M., Hanagaki,T.,

Hayatsu,N., Hiramoto,K., Hiraoka,T., Hirozane,T., Imotani,K., Ishii

,Y., Ito,M., Kawai,J., Kojima,Y., Konno,H., Kouda,M., Matsuyama,T.,

Nakamura,M., Nishi,K., Nomura,K., Numasaki,R., Okazaki,Y., Okido,T.,

Saito,R., Sakai,K., Sakazume,N., Sasaki,D., Sato,K.,

Shibata,K., Shingawa,A., Shiraki,T., Sogabe,Y., Suzuki,H., Tagawa

,A., Takahashi,F., Takaku-Akahira,S., Tanaka,T., Tomaru,A., Toya,T.

, Watabiki,A., Yasunishi,A., Muramatsu,M. and Hayashizaki,Y.

RIKEN Encyclopedia of Mouse Full-length cDNAs (Akimura,T., et al.

2001)

Unpublished (2001)

Contact: Yoshihide Hayashizaki

Laboratory for Genome Exploration Research Group, RIKEN Genomic

Sciences Center(GSC), Yokohama Institute

The Institute of Physical and Chemical Research (RIKEN)

1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan

Tel: 81-45-503-9222

Fax: 81-45-503-9216

Email: genome-res@gsc.riken.go.jp

URL:http://genome.gsc.riken.go.jp/

Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh

,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.

Normalization and subtraction of cap-trapper-selected cDNAs to

prepare full-length cDNA libraries for rapid discovery of new

genes. Genome Res. 10 (10), 1617-1630 (2000)

wagi,K., Fujiwaka,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E.,

Watabiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsura

,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and

Hayashizaki,Y.

RIKEN integrated sequence analysis (RISA) system--384-format

sequencing pipeline with 384 multicapillary sequencer. Genome Res.

10 (11), 1757-1771 (2000)

Konno,H., Fukunishi,Y., Shibata,K., Itoh,M., Carninci,P., Sugahara

,Y. and Hayashizaki,Y.

Computer-based methods for the mouse full-length cDNA
encyclopedia: real-time sequence clustering for construction of a
nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
Please visit our web site (<http://genome.gsc.riken.go.jp>) for
further details.

e mouse tissues.

FEATURES

source

Location/Qualifiers

1. .448

/organism="Mus musculus"

/strains="C57BL/6J"

/db_xref="taxon:10090"

/clone="G370007124"

/cell_lib="RIKEN full-length enriched, B16 F10Y cells"

/note="pooled tissues ; (tissue type=cerebellum,

dev stage=16 days neonate, sex=mixed),

(tissue type=cerebellum, dev stage=0 day neonate,

sex=mixed), (tissue type=hippocampus, dev stage=adult,

sex=male), (tissue type=whole body, dev stage=9 days

embryo, sex=mixed), (tissue type=lung, dev stage=13 days

embryo, sex=mixed)"

BASE COUNT 113 a 119 c 141 g 75 t

ORIGIN

Query Match 92.7%; Score 20.4; DB 10; Length 448;

Best Local Similarity 95.5%; Pred. No. 26;

Matches 21; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GAGCTTGACACGGCGGTTTCT 22

DB 420 GAGCTTGACACGGCGGTTTCT 399

Search completed: April 18, 2003, 07:32:04

Job time : 769.167 secs

